• Volume 5C, Number 4, June 1957

מוסד ויצמן למדעים במדעי הטבע והטכנולוגיה בישראל • ירושלים
The Weizmann Science Press of Israel • Jerusalem

BULLETIN OF THE RESEARCH COUNCIL OF ISRAEL

Section C TECHNOLOGY

Bull. Res. Counc. of Israel. C. Techn.

Incorporating the Scientific Publications of the Technion — Israel Institute of Technology, Haifa

BULLETIN OF THE RESEARCH COUNCIL OF ISRAEL

יוצא לאור ע"י

מוסד ויצמן לפרסומים במדעי חטבע ובטכנולוגיה בישראל

המועצה המדעית לישראל • משרד החנוך והתרבות • האוניברסיטה העברית בירושלים

הטכניון — מכון טכנולוגי לישראל • מכון ויצמן למדע • מוסד ביאליק

Published by

THE WEIZMANN SCIENCE PRESS OF ISRAEL

Research Council of Israel • Ministry of Education and Culture

The Hebrew University of Jerusalem • Technion—Israel Institute of Technology

The Weizmann Institute of Science • Bialik Institute

Manuscripts should be addressed: The Editor, The Weizmann Science Press of Israel, P.O.B. 801, Jerusalem

33, King George Ave., Telephone 62844

Volume 5C, Number 4, June 1957

מוסד ויצמן למדעים במדעי הטבע ובטכנולוגיה בישראל • ירושלים
The Weizmann Science Press of Israel • Jerusalem

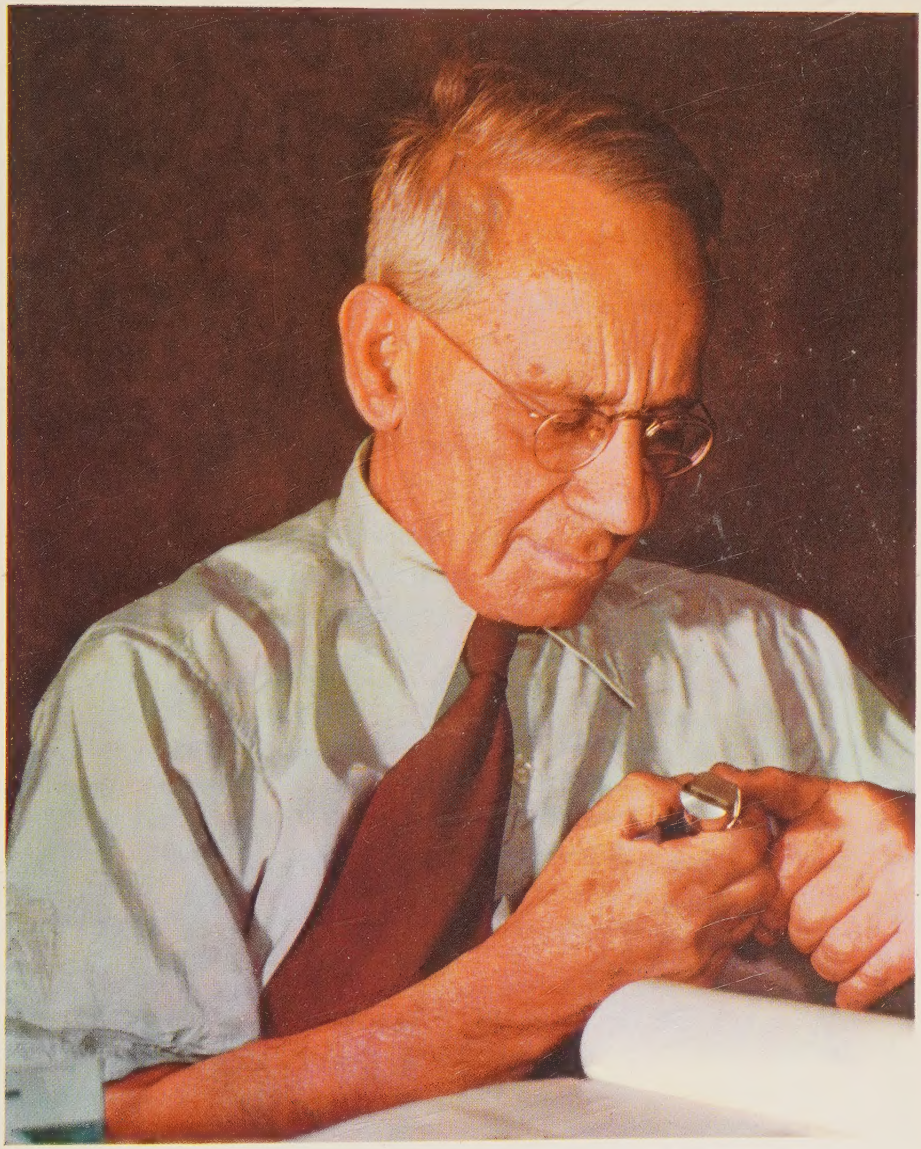
BULLETIN OF THE RESEARCH COUNCIL OF ISRAEL

Section C TECHNOLOGY

Bull. Res. Counc. of Israel. C. Techn.

Incorporating the Scientific Publications of the Technion — Israel Institute of Technology, Haifa

CONTENTS



PROF. EMANUEL GOLDBERG

"A chemist by learning, a physicist by calling and a mechanic by birth" — this definition of his profession is given by Prof. Emanuel Goldberg, to whom this issue is dedicated on the occasion of his 75th birthday.

Goldberg was born in Moscow in 1881 as son of a Medical Officer in the regular army, an extraordinary case in Jewish life at this period of Czarist Russia. His first scientific publication, devoted to electrochemical reactions in organic solvents, was made when he was 19 years old. The result led to a technical process patented by him in 1901 and exploited for many years afterwards by a prominent German company. This gave Goldberg the possibility of devoting his time to fertile scientific work beginning at the University of Moscow from which he graduated. He spent the years 1904 and 1905 in the famous laboratories of Walter Nernst in Goettingen and Wilhelm Ostwald in Leipzig, where the foundations of modern physical chemistry were being laid at the time. Goldberg obtained his doctorate at Leipzig University with summa cum laude honours. He dedicated the following two years to phototechnical problems while working as assistant to Prof. Miethe (then famous for his methods of colour photography) and as lecturer on photogrammetry at the Military Academy in Berlin.

In 1907 Goldberg was nominated as Professor at the *Akademie fuer graphische Kuenste und Buchgewerbe* in Leipzig. There he devoted his time fully to the science of photography and published his well-known method for casting neutral grey wedges, as an auxiliary tool for the densographs designed by him. This method found quite a number of applications in following years. At this time he also proposed a new method and built the apparatus necessary for the semi-automatic testing of the performance of photographic lenses.

He left Leipzig in 1917 to join the Board of Directors of *ICA A.G.*, one of the companies belonging to the Carl Zeiss Foundation. This company was later amalgamated with several others in the photographic field, resulting in the foundation of the *Zeiss-Ikon A.G.* Dresden, of which Goldberg was the first General Manager.

At the same time he also taught at the Technische Hochschule, Dresden and in 1923 published his, now famous, book *Der Aufbau des photographischen Bildes*, which was probably the first modern approach to the evaluation of the basic elements of modern photographic reproduction. The new concepts introduced in this book aroused considerable interest, especially since Goldberg described therein many new methods and measuring instruments. Whereas the previous concept was that the goal of photography is the true reproduction of nature, Goldberg's investigations proved (now it seems unbelievable that this had to be proved!) that there is a basic

difference between pictorial and applied photography. In pictorial photography, not the true rendering of light values, but only the impression on the person seeing the picture is of importance. On the other hand many applications of photography as for instance sound photography are possible only when a true rendering of light values is assured. With the conception of a law for the best transcription of this value (known as the "Goldberg Condition" — $\text{Gamma}_{(\text{negative})} \times \text{Gamma}_{(\text{positive})} = 1$) one of the theoretical cornerstones of modern sound cinematography was laid. — Of the numerous other methods described in the book, a process for measuring the speed of photographic emulsions may be mentioned which was later adopted as DIN Standard.

Although Goldberg was at this period primarily preoccupied with the expansion of the *Zeiss Ikon* concern, he still found time to continue to be the living spirit behind the technical development of the company, being responsible for the design of numerous instruments and cameras of which the *Contax* Camera may be mentioned.

His activities in Germany were interrupted suddenly with the advent to power of the Nazis. He was forced to leave Germany but his children were permitted to join him only after he consented to continue to work for the *Zeiss Company*. From 1933 to 1937 he was manager of the companies owned by the *Carl Zeiss* concern in Paris. He himself speaks of those years as having been spent in a golden cage. Only his decision to emigrate to Palestine in 1937 restored to him his personal and intellectual freedom.

In Tel-Aviv he established a small laboratory for applied optics and with the advent of the war, was made scientific adviser to the Allied Forces in the Middle and Near East. In this capacity he visited India for several months and helped there to set up a small scale optical industry. When the American Army arrived in the Middle East they also availed themselves of his experience in the optical field and assisted him in considerably expanding his laboratory. He is now adviser to the Israel Army Forces and Ministry of Defence.

His aims of laying the cornerstone to an optical industry in this country and of educating a generation of young specialists, were furthered when in 1947 the *Goldberg Instruments Ltd.* company was formed, whose General Manager he has been since its foundation. This company specializes in the development and production of military instruments and in the manufacture of optical instruments designed by Goldberg, e.g. a line of industrial refractometers.

Goldberg's thoughts were always dominated by an idée fixe that the culmination point of photography will be met only when its possibilities for the organization of human mental products are fully exploited. In this direction many advances were made by him, beginning in 1926 when he made a world record for the smallest copy of a document (a fully legible page on 0.01 mm^2), a record unsurpassed even now. In 1931, at the occasion of the 8th International Congress for Photography, he demonstrated his "Statistical Machine", a device where for the first time electronic

means were used to search from out of a series of photographic images on a long "microfilm" roll, those images which are of specific interest to the user. The work on this subject, far in advance of the development of bibliography at that time, was unfortunately suddenly interrupted with the advent of Hitler, and Goldberg, in his new surroundings and later with a war going on, had to interrupt his work in this direction. It should be mentioned here, that when Vannevar Bush (then Director of the U.S. Office of Scientific Research) envisaged in 1945 in an article in *Life* magazine a similar device, termed "The Rapid Selector", the true inventor, who had actually built and demonstrated such a machine years before, was not mentioned. After the war, Goldberg again turned his attention to the vast possibilities of photography as a tool for documentation and designed a radically new microfilm camera, the manufacturing rights of which were taken over by a company in the United States.

Now, after his 75th birthday, his dream is still as it was thirty years ago: a library consisting of an automatic "record" changer with viewing screen, equipped with discs bearing minute photographic images made according to his process and holding not less than 25,000 books of 250 pages each. With 10 discs such a "library" would contain the equivalent of a quarter of a million books, and would occupy less space than a conventional writing desk. His friends and pupils wish him many more healthy and fruitful years to see these dreams come to reality.

No better tribute to his many achievements could have been made than a renewal of his Doctor's Diploma, sent to him some months ago by the University of Leipzig, reading as follows:

> "Die Mathematisch-Naturwissenschaftliche Fakultaet der Karl Marx Universitaet Leipzig erneuert Herrn Professor Dr. Emanuel Goldberg das ihm am 21. April 1906 ausgestellte Doktor-Diplom und spricht dem Jubilar fuer seine Verdienste um die Klaerung grundsaetzlicher Fragen in der wissenschaftlichen und angewandten Photographie zu seinem goldenen Doktor-Jubilaeum die herzlichsten Glueckwuensche aus."

S. NEUMANN

ON THE QUALITY OF PHOTOGRAPHIC REPRODUCTIONS

A. Narath

Institute for Applied Photochemistry and Film Technics, Technical University, Berlin

ABSTRACT

It appears that the problem of reproducing "details of sharpness" (Goldberg) has still to be thoroughly investigated. It is closely related to the problem of reproducing details of brightness. The latter has been thoroughly investigated by Goldberg, who published the results in his book on the structure of the photographic image. It is the merit of Higgins and Jones to have continued the investigations, so that today we can survey a great number of factors which have to be taken into account. Regarding the rendering of details of sharpness, it is to be hoped that future investigations will heed the basic and thoroughly elaborated points of view established in this paper, namely, that only by interrelating both problems and by simultaneously considering the reproduction of the details of both the brightness and the sharpness, can a satisfactory solution be expected and the aim be reached which was in Goldberg's mind from the beginning.

When Goldberg more than 30 years ago published his fundamental work on the structure of the photographic image[1], the experts were for the first time given a comprehensive presentation of one of the most important problems of photography, which even today is still topical, and will probably always remain so. Starting with theoretical considerations, Goldberg created over the years the assured experimental basis he needed in order to establish harmony between theory and practice. The first precondition was the development of suitable and dependable methods and apparatus for measurements which were required in carrying out extensive examinations. The Goldberg Detail-Plate, Goldberg Wedge and Goldberg Densograph (Figure 1) were the results of his creative activity; these are forever attached to his name and have rendered great services to many research workers during their scientific undertakings.

The difficulties of the problems attacked by Goldberg can be seen best from the fact that even today they have not yet been definitely solved, although much scientific research has been done over the past decades in these fields. An attempt will be made below to give a short survey of the causes responsible. It will be seen that it is necessary to consider not only details of brightness but those of sharpness as well, a problem with which Goldberg also occupied himself and which he wanted to treat in the second part of his work on the structure of the photographic image.

Received July 2, 1956.

I. REPRODUCTION OF DETAILS OF BRIGHTNESS

(*a*) *Special conditions for rendering of object brightness in pictorial photography*

The entire process of the pictorial reproduction of object brightness can be split into an objective and a subjective phase. In regarding an object, the eye is replaced by the photographic camera. The camera has to record everything seen by the eye. It has to be stipulated that the impression gained by viewing the positive, either as diapositive or as paper print, be as near to identical as possible with a direct viewing of the object to be photographed. The object is characterized by a distinctive distribution of the luminances of its individual elements, and it has a characteristic range of luminance and a mean luminance. In using the elementary values of light in the whole process of photographic reproduction, it has to be remembered that these values apply solely to the V_λ curve of the eye adapted to brightness; therefore, the spectral composition of the light emanating from the object, as well as the spectral transparency of the transmitting elements (objectives, glasses, filters, etc.), and the spectral sensitivity of the photographic layer, and, lastly, the spectral composition of the light on regarding the positive, have to be considered. Luminance and its distribution can be measured objectively on the object to be photographed, and the most important value is the "real object range" (ROR), which, according to Goldberg, is defined as the logarithm of the proportion of greatest to smallest luminance in the object. The object range is diminished by the so-called "air-light" between object and observer or camera, respectively, and further by reflected and dispersed light in objective and camera, and is then called "usable objective range" (UOR). In contrast to this is the "subjective object range" (SOR), which, according to Goldberg, is the ROR reduced by the "air-light". As the photographic process, in general, has only the task to reproduce so that on viewing the positive the same impression is received as on regarding the object to be photographed, the SOR only has to be rendered.

The objective phase of the whole reproduction now includes all the conditions which have to be fulfilled in order to render physically correct any given object range, whilst the subjective phase includes the eye in the considerations and may either render more difficult or facilitate the conditions to be met. It is Goldberg's special merit that in all his investigations he put the subjective phase in the centre of his observations and thus deduced the conditions necessary for a reproduction true to nature.

Considering first the objective phase, a reproduction true to nature has to be made of both the mean luminance of the object and the range of luminance. The distribution of luminance of the object corresponds to the distribution of the intensities of illumination in the ground-glass plane, which, apart from losses through absorption and reflection in objective and camera, are but a function of the luminance B of the individual points of the object and the relative aperture of the objective.

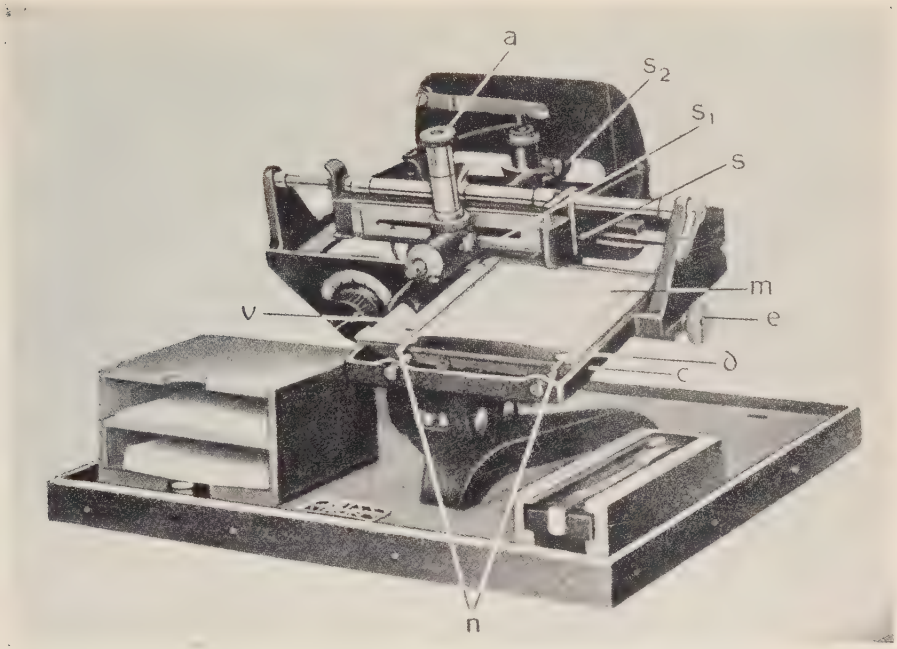
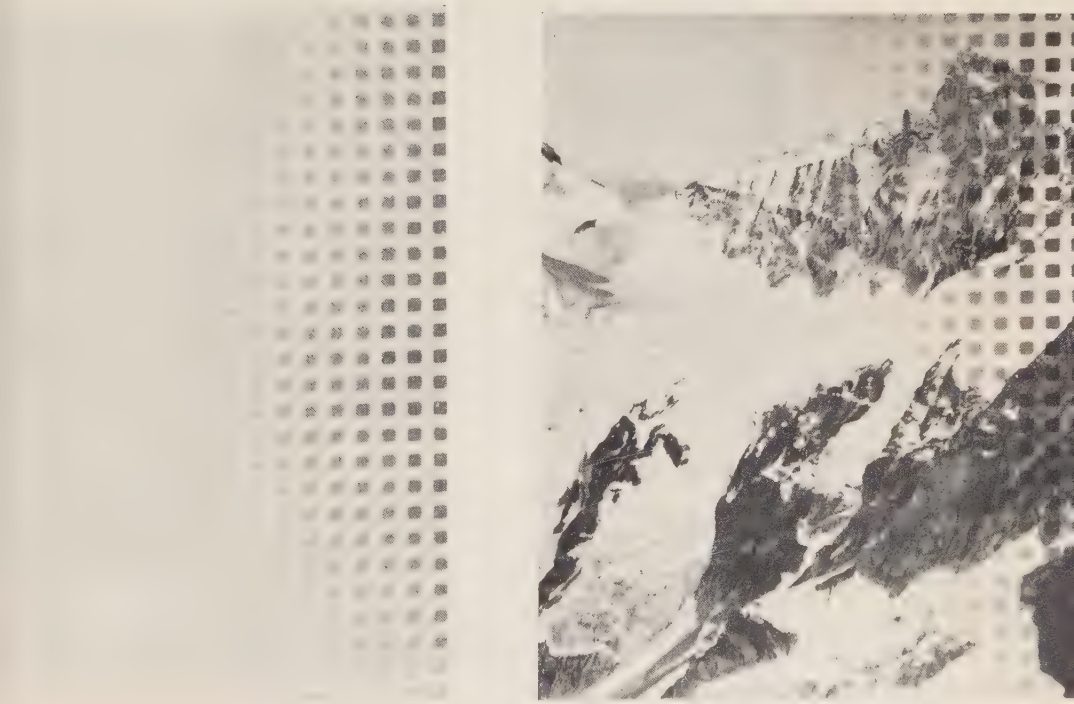
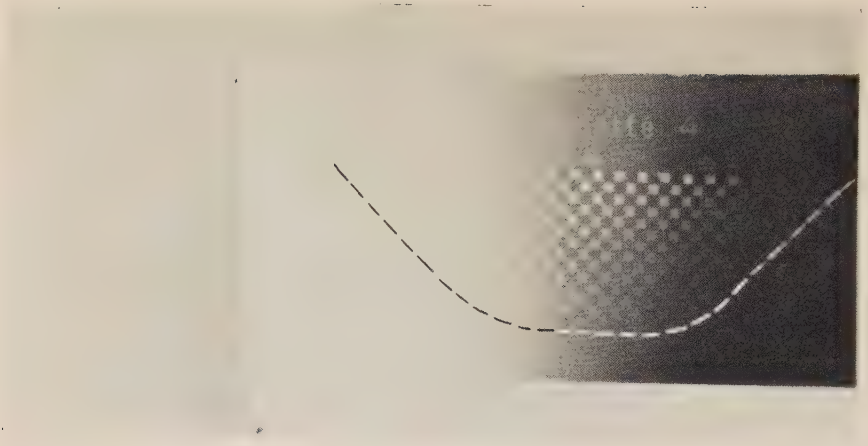


Figure 1
Goldberg Densograph.

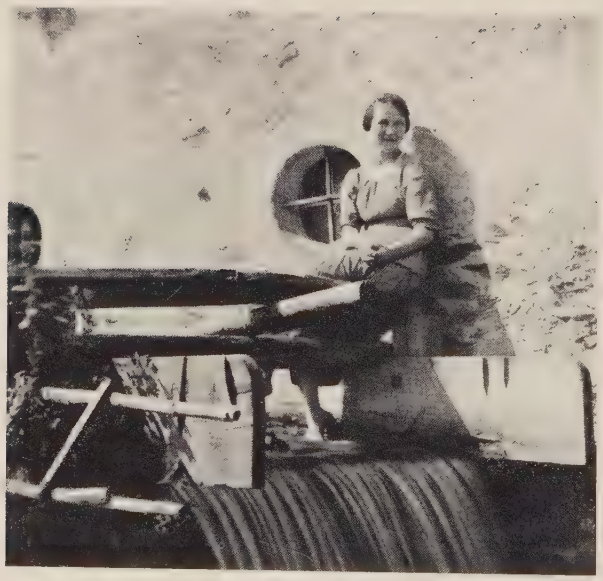


(a) Figure 2 (b)

ermination of rendering of details from photographed objects according to Goldberg, by means of



(a)



(b)

Figure 3

Determination of rendering of details of photographic layers, according to Goldberg. (a) Detail-plate printed on photographic paper with detail curve. (b) A negative printed on the same paper.

For a physically correct reproduction, the transparency of the positive T_P or the remission of the positive R_P, respectively, must be proportional to the exposure of the negative E_N and thus also to the intensity of illumination in the ground-glass plane, as long as the exposure time remains constant. Given D_N and D_P as the densities of the negative and the positive, respectively, then the condition $(dD_N/d\log E_N)\cdot(dD_P/d\log E_P)=1$ has to be fulfilled, or the condition $\gamma_N\cdot\gamma_P=1$, which applies only to the straight parts of the density curves. A process in which the second condition is fulfilled is called "straight line recording", whilst the fulfilment of the first condition is called "toe recording"[2]. Both these conditions were first derived by Porter and Slade[3], who were also the first to establish a method for constructing from a given negative curve that positive curve which ensures a rendering true to nature. This problem had previously been dealt with by Hurter and Driffield[4]. The latter postulated in their basic works the relation $D_N=\gamma_N\log(E_N/i_N)$ for the density curve, in which i stands for the inertia and from which derives $T_N=i_N^{\gamma}/E_N^{\gamma}$. They attempted to set up a condition for the "perfect negative" and put for this purpose the equation $T_N=1/E_N$ at the beginning of their work. Elsewhere in their work they said that for a "perfect negative" $\gamma_N=1$ must be chosen. With the exception of the constant i_N, the latter condition is identical with their demand for a "perfect negative". It can be seen that Hurter and Driffield had not yet found the correct condition of printing; it appears from their assertion that the sum D_N+D_P must be constant, which ensures a reproduction true to nature only for $\gamma_N=\gamma_P=1$, but which is wrong for other values of γ; this fact was pointed out to them especially by Porter and Slade.

Hurter and Driffield started their considerations from the characteristic curve drawn by them and their remark that only the straight part is the "period of correct representation"; this already shows that they did not foresee the possibility that curved parts, too, allow for a correct reproduction, if these are compensated by corresponding bends of the positive density curve. The first to recognize this were Porter and Slade, who also pointed out that with practically given density curves only small light-ranges are reproduced true to nature. These facts in no way diminish the great merits of Hurter and Driffield for the inception of sensitometry, especially when it is recalled that their basic works are dated 1890 and 1891.

In the objective phase of the entire process of reproduction it must furthermore be considered that the layers possess a grainy structure and the measured values of transparency and density depend on the manner of transillumination. This means a consideration of the Callier-effect, which is important in printing and in projecting the positive. Lastly, the wavelength dependence of the transilluminated layers has to be considered, because of selective absorption and scatter in both silver grains and binder.

When all the above factors are known, the copy curve can be drawn, which is the combination of the density curves of negative and positive. $\gamma_N\cdot\gamma_P=\gamma_C$ applies to the straight part of this so-called copy density curve. The 4-Block method indicated

by Jones[5] can also be used, where the luminance of the reproduction of a positive can be compared to the luminance of an object on its being photographed. Theoretically, $\gamma_C=1$ and the curve should be straight, i.e. without curvature at $D=0$, and should also be sufficiently extended in order to cover the entire object range reproduced. This purely physical demand is the first problem of the objective phase of the whole process. It is found that these ideal conditions cannot be realized in practice and the question arises whether deviations are permissible and to what extent. This leads to the second problem, which can be solved only by considering the properties of the eye and which includes the subjective phase of the whole process. There are two alternatives: either the physiological functions of the eye are considered, or experiments with various, suitably chosen objects to be photographed will have to clarify which conditions have to be met and what deviations are permissible from a physical point of view. Realizing the importance of this posing of the problem, Goldberg paid special attention to the investigation of the subjective phase of the whole process, discussed thoroughly both the above-mentioned alternatives for arriving at a solution, and carried out extensive experiments.

In order to solve the first problem, the object range to be rendered has to be known. The ROR can have extraordinarily high values and even the SOR has luminances ranging from 10^{-8}—10^{5} stilb*. In the following we shall confine ourselves to the SOR, as only the latter needs to be rendered when nature subjects are photographed. But it has to be taken into consideration that this range, which covers many powers of ten, also includes the adaptation of the eye to light and darkness. It may be assumed that the eye is adapted to light whilst looking at a positive. The transition from adaptation to darkness to adaptation to light occurs at luminances in the range 10^{-2}—10^{2} stilb. Taking as a base the lower limit, the object range would still be 10^{7}. Obviously, substantially smaller values may be expected in practice, because there is no need to use extreme apertures. However, in special cases such as doorways, object ranges of 10^{6}—10^{7} may occur. Goldberg found the numerical values of object ranges correlating to a great variety of objects. A reproduction true to nature demands the fulfilment of $\gamma_C=1$ for the copy curve. Theoretically, this curve should rise in a straight line from $D_C=0$ to a value D_{max} which should be such that the particular object range could be reproduced in its entirety. These rigid demands cannot be realized in practice. Here we have to obtain larger density ranges on a diapositive than on a paper print. However, the object range is liable to be appreciably reduced by projecting a diapositive onto a screen by reason of the latter's properties and of scattered light in the room. The density range of paper prints is considerably smaller and is limited by the albedo which at best allows for a value of 1.7 for the darkest spots. Here, the kind of lighting and the manner of viewing the image are extremely important. The density curves of printing papers (with the exception of pigment papers) generally have a very limited luminance. These curves do not rise in a straight line from zero, but possess a more or less pronounced curvature.

* 1 stilb = 1 cd/cm².

This part of the curve is of special importance in the rendering of lights. It can therefore be predicted that the demands made for reproduction true to nature of object ranges cannot be met. Copy curves attainable in practice can be arrived at by combinations of different density curves of the materials used for negatives and positives; but it will be found that the ideal curves deviate more or less strongly. The problem is to find permissible deviations and the conditions for a satisfactory rendering of the photographed object.

It is at this point that the second problem comes in which concerns the subjective phase. Undoubtedly, it will not be necessary to render the entire object range, which in its extreme case, according to Koenig, covers 660 grades of brightness discernible by the eye. Goldberg called these grades "details of luminance" and means by it the logarithm of the ratio of two barely discernible brightnesses. He found from extensive researches that the values of $Dt=0.02$ for lights, $Dt=0.04$ for medium tones and $Dt=0.1$ for shadows, lead to satisfactory results. Furthermore, he also found that the magnitude of these "minimum details" largely depends on the nature of the object and that smallest structures lead to an enlargement of these details. This influence appears clearly in the "tone-curves" constructed by Goldberg for a variety of objects. When the density curve is differentiated, then the differential curve should afford an objective statement as to the quality of the reproduction of details; when $\gamma_C=1$, the copy density curve should give a rendering true to nature of all details of brightness of the photographed object. Experiments showed, however, that with the inclusion of the eye, the detail curves possess a form entirely different from that to be expected from the differential curves. For this reason Goldberg evolved his well-known "detail-plate" and with its aid found the detail curves. Then it was found that even with $\gamma=1$, there is a certain loss of detail. Goldberg attributes this loss of detail mainly to the distribution of the silver grain in the layer, which is connected with the resolving power, the granularity of the paper, etc. Here, too, it is absolutely clear that in judging photographic images the eye cannot be ignored. Goldberg's success in obtaining, with the aid of the "detail-plate", measurements commensurate to the subjective conditions of observation, has therefore to be considered an especially great achievement. Goldberg examined the rendering of details of photographic objects by inserting a "detail-plate" into a telescope (Figure 2) and that of photographic layers by contact printing his "detail-plate" on printing materials (Figure 3).

There is also an alternative method for considering the rendering of details and conditions of observation. Jones[6] improved his 4-Block method so as to take into account primarily all physical factors of the reproduction process and thus arrives at an illustration with 8 Blocks (Figure 4). In Block 8 the luminances in viewing diapositives or paper prints are shown as functions of the luminances of the object and the resulting curve shows the ensuing deviations. Jones completed the 8-Block method by adding three more Blocks in order to take care also of the subjective phase of reproduction. Curves showing the dependence of the observed intervals

on the subjective luminance are needed here. Jones found these curves experimentally for a test area covering an angle of view of 2° and the two halves of which had been subjected to changes in their luminances so that the luminance differences could be just discerned. These curves were measured for various mean luminances as the latter differ, of course, considerably in viewing the object, the diapositive and the print respectively. Therefore, a curve corresponding to 20 ft L. was drawn in Block 9, and one for 660 ft L. in Block 10. Block 11 then shows the working of the reproduction, in grades of brightness, measured for the entire process (Figure 4). Unfortunately,

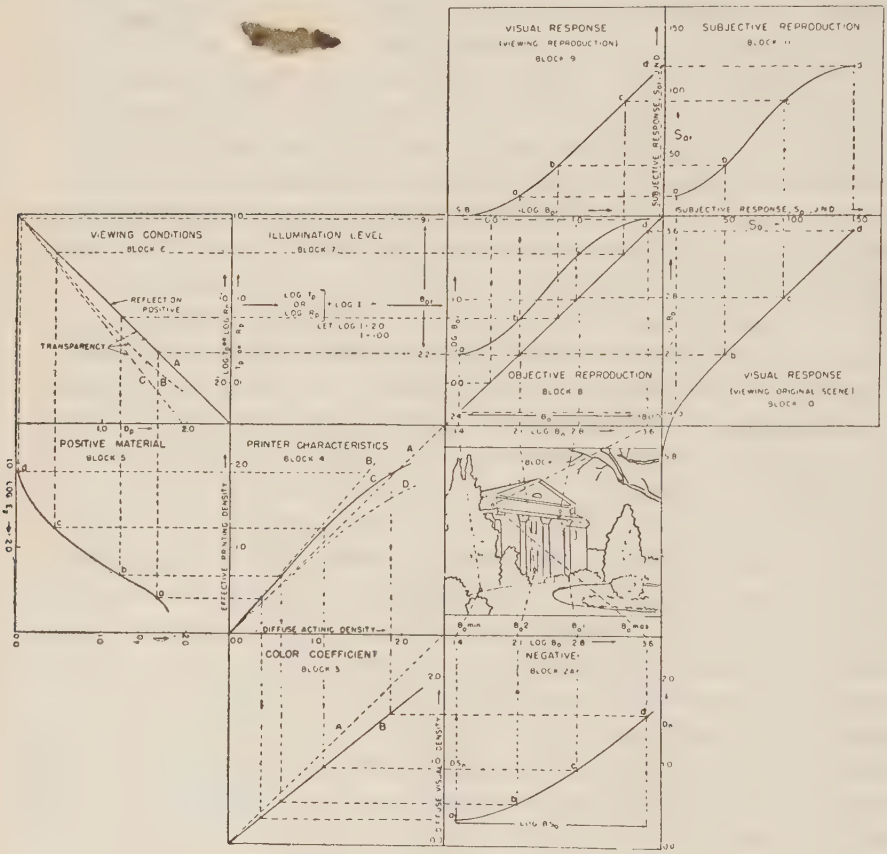


Figure 4

11-Block method according to Jones for the determination of rendering of luminance of photographed objects.

it was found that the resulting final curves did not correspond exactly to practice, although care was taken to consider all conceivable factors. The chief argument against this method is the fact that the evaluation of the curves representing the

threshold values of the luminance does not coincide with the conditions of observation in practice. Goldberg already pointed out that the nature of the object plays an important role in finding the minimum detail and that smallest structures greatly influence the size of these details. Jones admitted that his method would rarely lead to binding results and he finally chose a process for finding optimum printing conditions which to the present day is the most certain and which was also used by Goldberg in his extensive researches. This method consists in producing as great a number of negatives as possible of the most varied objects, printing them on different materials and having them judged by a large number of viewers. This is a

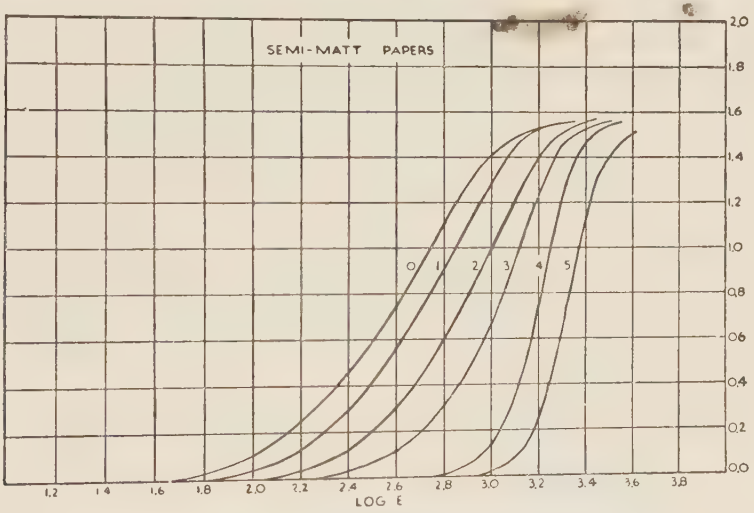


Figure 5
Density curves of semi-matt papers of varying contrast.

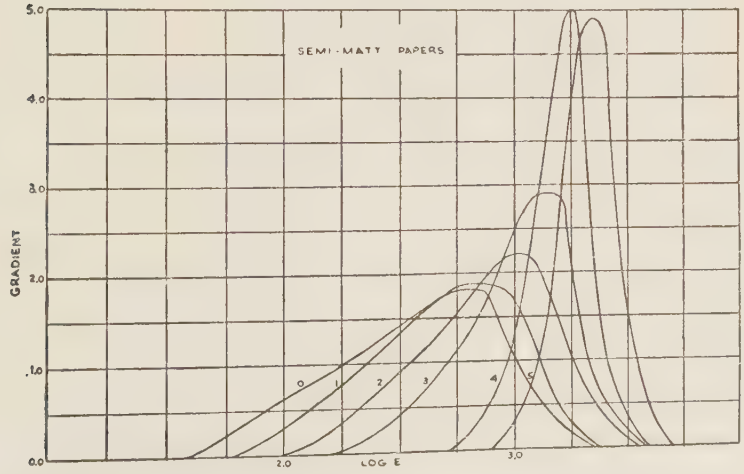


Figure 6
Differential curves of the semi-matt papers.

statistical method ensuring without doubt the greatest possible success. Comparing the results thus obtained with sensitometric data leads to interesting deductions on the execution of the entire process of reproduction. Jones examined 170 negatives and printed each on 6 photographic papers of different contrast (Figures 5 and 6). The viewers sorted the prints according to their quality as first, second and third choice. In addition, all positives were evaluated sensitometrically and the density curves of the positives showed the areas on the negatives containing maximum and minimum densities, i.e. the position of the density range of each negative on the density curve of the positive material. This most painstaking work produced very interesting results. In the following we shall not discuss each individual curve and it is also impossible to make comparisons with the numerous previously published curves; but a few examples will show the results obtained by Jones and the deductions made for practice and theory of the process of photographic reproduction.

The density ranges of the 170 negatives showed an almost Gaussian distribution with a mean value of $\Delta D=1.25$. The maximum values did not exceed 2.5. Another 1000 negatives, taken by amateurs and developed in various photographic shops, had a very similar distribution with a mean value of $\Delta D=1.30$. The mean value of the range of luminance of the evaluated 126 outdoor pictures was approximately 160 or $\Delta \log B=2.2$, which is higher than that hitherto taken for the average object ($\Delta \log B = 1.5$). This value results, according to Jones, from the elimination of stray light in the photometers used.

All of the negatives having been printed on paper, it was possible to find the positive density curve with which each of these negatives gave a "first choice" print. Figures 7 and 8 illustrate the curves for a soft and a normal paper, respectively, and also show the density ranges of the various negatives. The results are most interesting and, at first sight, unexpected. The density curve of the positive is utilized only up to the beginning of the shoulder, to an average of $D_{P\max}=1.4$; however, — and this is most astonishing — the light-ranges by far transcend the toe where there is no more blackening at all. The dispersal of values is here particularly wide. The first conclusion to be drawn from these results is that for a correct exposure of the positive the value of $D_{N\min}$ is of primary importance, but not that of $D_{N\max}$ or even the mean value of D_N; this makes possible the foundation of a sensitometric method for the evaluation of density curves of printing papers. The second conclusion to be drawn from these results is still more significant. On the average, all of the photographs use only the toe and the adjoining straight part of the positive density curve, thus never exceeding the shoulder. This means that the rendering of details in the shadows is not renounced. The opposite applies to the lights. Here, the range of luminance not only enters the toe but sometimes goes far beyond it and one might be tempted to conclude that the lights may be without details. This would be, however, a rash conclusion. It must not be overlooked that the dispensation with details in the lights exclusively depends on the object to be rendered. As can be seen from Figures 7 and 8, the brightest spots in the positives lie throughout in a

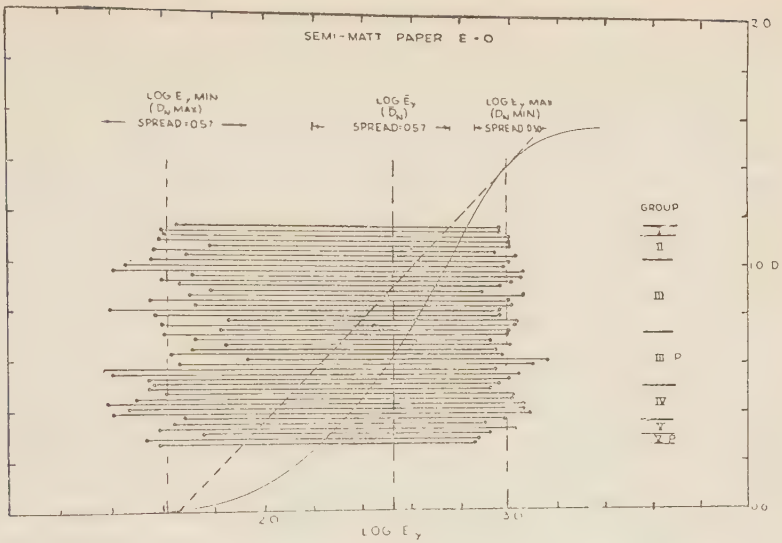


Figure 7
Position of density ranges of 45 negatives on an extra-soft paper, according to investigations made by Jones.

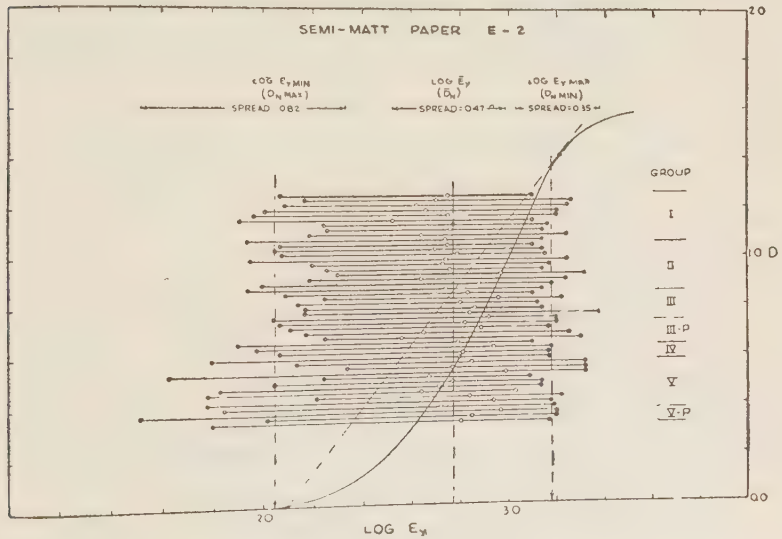


Figure 8
Position of density ranges of 47 negatives on an approximately normal paper, according to investigations made by Jones.

density area of the copy curve, which allows for a satisfactory rendering of detail. From the fact that the statistical mean value lies at the basis of the curve or even outside it, we can draw no conclusion other than that perfect white is generally wanted at the brightest spots. But it will depend entirely on the kind and size of these

spots whether the rendering of details may be renounced or not. The conditions in the shadowy parts are somewhat dissimilar. It could have been expected that a certain lack of detail was to be taken into account. However, experiments showed the dispersal of the values to be small in proportion, and the light-range practically never extends into the shoulder of the copy curve; this obviously means that good rendering of detail is wanted in the shadowy areas. In this case, too, no hasty conclusions should be drawn, as besides the darkest areas of the object the half-tones should also be satisfactorily rendered. Placing the shadows too high in the upper part of the positive curve may render the half-tones lacking detail, which would run counter to a satisfactory rendering. Everything therefore depends on the grades of tones occurring in the image and which of them are important for it. It appears from the aforesaid that the diagnoses of Jones and Goldberg need not contradict each other. The prevalence of details in both lights and shadows will have to be evaluated as to the nature of the object. This evaluation will also depend on the personality of the viewer, on regional differences, i.e. land, climate and way of life of the inhabitants, and, last but not least, on fashion.

One thing can be said with certainty: as the toe and the straight part of the positive curve are always used simultaneously, it is impossible to fulfil the condition $(dD_N/d\log E_N)\cdot(dD_P/d\log E_P)=1$ for the entire object range. Jones believes that fulfilling this condition, at least for the medium tones, guarantees a "first choice" print; but from the few examples he analyses and publishes, it appears that generally the product is greater than one. There occur considerable curvatures in both the lights and the shadows, i.e. compressions in the gradation of light, which are inevitably due to the form of the positive and negative density curves in their upper and lower regions. If—according to Jones—the product of the gradients must in some cases exceed 1, this only confirms the necessary deviation — probably first discovered by Goldberg — from the physically required condition, so as to achieve an objectively correct reproduction. Goldberg was not in a position to give a perfect explanation for the ensuing loss of detail, but he assumed this loss to be due, besides other causes, to the distribution of the silver grain in the layer. This problem will be treated later when it will be shown that reproduction true to nature of luminances of individual portions of the image also depends on their area; therefore, an important factor of the objective phase has been ignored in the considerations made so far.

(*b*) *The general condition for the rendering of object brightness*

In order to ensure a reproduction true to nature in pictorial photography, the transparency of the positive should be proportional to the exposure of the negative, when all subjective factors are excluded, thus $T_P=b\cdot E_N$. This nessecitated the application of the γ-method. In reproducing light-ranges, this is only a special case of a more general condition expressed by

$$T=a+b\cdot E,$$

and which asks for a linear connection between transparency and exposure. This formulation also includes the case of linear connection between transparency and exposure of the negative; then $b<0$. With $b>0$, we find the following three cases: $a<0$, $a=0$ and $a>0$. These different conditions lead to various processes important in sound film techniques. In this application of photography there is no need for a proportionality between transparency and exposure as long as a linear connection between these two values exists. Considering now the reproduction of a brightness-range in the negative-positive process, then the transparency curve of the copy has to fulfil the equation $T_P=a+b\cdot E_N$. Replacing the transparency with the density and the exposure with its logarithm, results in

$$D_P = 1/(a+b\cdot \exp_{10}\log E_N).$$

For $a<0$ and $a>0$ this leads to two different copy density curves, which run concave and convex to the abscissa, respectively. It is easily seen that the form of both these curves is independent of the absolute values of the constants a and b. The rise of these curves is not constant as in the γ-method, but varies continuously with $\log E$. The author[2,7] called these the α- and β-methods. Finally, by taking $b<0$, a connection between transparency and exposure of the negative is obtained and we can arrive at the usual demonstration by substituting density for transparency and the logarithm of exposure for exposure. Then we obtain that form of the negative density curve which alone makes possible a linear connection between T_N and E_N. This was called the δ-method by the author. All these methods are represented in Figure 9, including

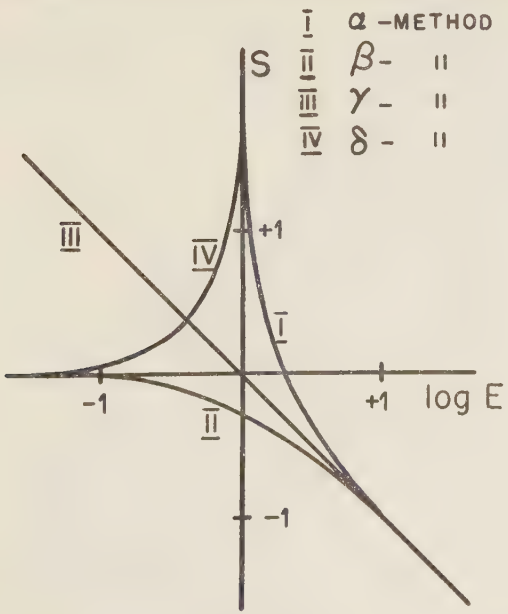


Figure 9
Illustration of density curves of linear functions of transparency, according to Narath. Negative-positive process: α-,β-, γ-method. Negative process: δ-method.

the γ-method which results in a straight line inclined at an angle of 135° to the abscissa. The various methods are listed in Table I, the last column of which shows the magnitude of steepness; the latter changes continuously in all but the γ-method where it remains constant.

TABLE I

Method	Equation	$dD/d\log E$
α	$D_W = -\log(-a + bE_A)$	$1 - \infty$
β	$D_W = -\log(a + bE_A)$	$0 - 1$
γ	$D_W = -\log b - \log E_A$	-1
δ	$D_W = -\log(a - bE_A)$	$0 - \infty$

The curves in Figure 9 may be moved parallel to either the abscissa or the ordinate; parts of curves measured in practice can thus always be examined as to whether they coincide with particular parts of the theoretical curve. There exists a linear correlation between transparency and exposure in this range. In conclusion the following rules result:

(i) There exists only one form of the density curve having a linear connection between transparency and exposure of the same negative material.

(ii) There exist three different forms of copy density curves having a linear connection between transparency of the copy and exposure of the negative.

In pictorial photography, however, only in applying the γ-method is there a correct reproduction of brightnesses possible, because here we have to demand proportionality between transparency of the copy and exposure of the negative. The conditions in sound photography differ from pictorial photography, as in the first case the

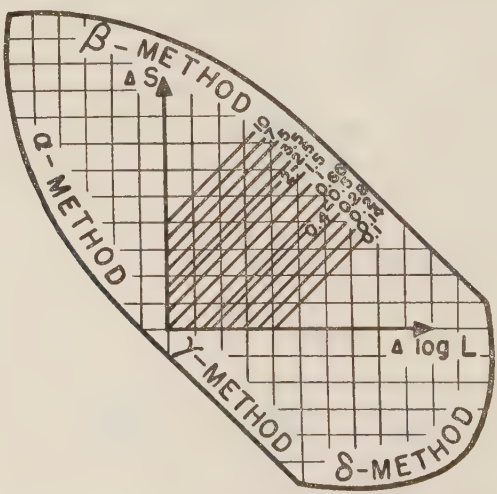


Figure 10
Chart for the determination of density curves of linear functions of transparency and the simultaneous determination of transfer factors, according to Pistor.

brightness of the recording light oscillates around an average value of light, according to the magnitude of the amplitudes of changes. The conditions prevalent here resemble that of an electronic valve diagram. By taking a known bias, a certain average value is obtained, on which the amplitudes of the alternating current will be superimposed. Pistor[8] developed a gauge for the aforementioned general density curves of linear transparency functions as shown in Figure 10 and which, when laid over a given curve, enables one to state quickly which part corresponds to the methods as shown in Table I.

II. REPRODUCTION OF DETAILS OF DEFINITION

An object, on being projected into the ground-glass plane of the camera, gives a distribution of brightnesses which has to be reproduced true to nature by the photographic process. One may think of the whole area of the image being divided into a series of small squares, each of them having a constant intensity of illumination. When these elements are sufficiently small, a continuous rise of the intensity of illumination can be replaced by this discontinuous curve. Although abrupt changes in the brightness of neighbouring areas of the image are prevented by the refraction of light, one can still notice more or less defined jumps in brightness, which Goldberg termed "details of brightness". For these, reproduction true to nature has to be required and the points of view relating thereto have already been fully discussed. But the conditions for a correct reproduction of the contours of adjoining elements of the image differing in brightness have yet to be found. For this Goldberg coined the term "details of sharpness". It can now be seen that the demand for a reproduction true to nature of all the contours appearing in the object is only a special case of a more general demand — the geometrically true reproduction of all areas of the object. The difference between these two formulations becomes apparent when it is recalled that an area may sometimes be reproduced in different sizes — caused, for instance, by the diffusion halo in the photographic layer — and still possess identical sharpness of contours in both cases. The author therefore proposed the name "fidelity of area"[9]. An element of area can therefore be reproduced with sharp contours either widened or narrowed; it may also happen, however, that the contours are not sharp, i.e. a zone of transition of diminishing density arises which also prevents a geometrically true reproduction. When two elements of areas with unsharp contours are so near that their zones of transition overlap, any further approximation may cause both areas to flow into each other without interval and they are no longer "resolved". The "resolving power" and the "sharpness of contours" are therefore two special cases of the more general concept of the "fidelity of area".

A ground-glass image may vary considerably in its content of area details. The contents are less than those of the object to be photographed, as is the case with details of brightness, due to "air-light" and the properties of the objective, particularly its aberrations. The amount of area details obviously depends strongly on the object. This special problem has recently gained particular prominence in television in

connection with the elimination of redundance for the purpose of narrowing frequency bands. Cherry and Gouriet[10] defined the details in a line l_2—l_1 with

$$\mathrm{Dt} = \int_{l_1}^{l_2} \left| dU/dx \right| dx,$$

U being the amplitude and x the line-coordinate. For natural objects and images on films they obtained values from 1.8% to a maximum of 8% of all area elements ("points" in television) of an image area fully modulated with 3 MHz on a chessboard pattern. The amount of area details of varying brightness is, therefore, mostly very small.

In the reproduction of area details, too, we have to distinguish between a subjective and an objective phase. The properties of the human eye have to be taken into account in the subjective phase, The "sharpness of vision" is limited by the retina consisting of individual elements sensitive to light, and is of a magnitude of approximately 1′, which defines the distance between two lines being barely seen separately. The sharpness of vision thus corresponds to the photographic resolving power. On the other hand, the ability of the eye to perceive a single line as such depends only on the volume of light exciting the retina; such a line may therefore be considerably smaller than the diameter of a cone in the retina. Something quite similar applies to photography. Individual lines of very diminutive diameter can be reproduced, whilst more such lines in close proximity fuse into each other in certain circumstances. The possibility of recording such a thin line is again a problem of the intensity of illumination, sensitivity of the photographic layer and, as will be shown later, the magnitude of scattering of light in the layer. Apart from effects such as the Eberhard, the Ross and the Kostinsky effects[5], which may be largely prevented, the fidelity of area is impaired in the first place by the scattering of light in the photographic layer and in addition by the graininess of the image. The sharpness of contours will always play an important part in area elements, independently of the size of the latter. When two large areas are separated by a narrow margin so that the transition zones of the haloes of diffusion overlap, then the margin may disappear. This property of photographic layers is naturally a consequence of contours lacking sharpness; it is called "resolving power" for a given test object consisting of lines separated by intervals (see Frieser[11]). The magnitude of the resolving power obviously depends on the line/space ratio and it is usually given the value of 1. The influence of the optics, in the reproduction of line gratings, on the resolving power, particularly due to effects of refraction, and influence of the aperture of the objective in connection with the thickness of the photographic layer, were thoroughly investigated by Narath and Schimmel[12]; furthermore, the hitherto proposed test objects were compared with each other both theoretically and experimentally and it was attempted to deduce a general function of turbidity.

It would be interesting to know, from the theoretical point of view, whether, by exposure with a given distribution of power of illumination, the resulting densities

and the found distribution of the effective illumination afford the prediction of expected densities for any other distribution. As in Figure 11, the "step function", "rectangular pulse", "rectangular curve" or "sinusoidal curve" may be chosen. Using, for instance, the "step function", one could find the effective distribution of light

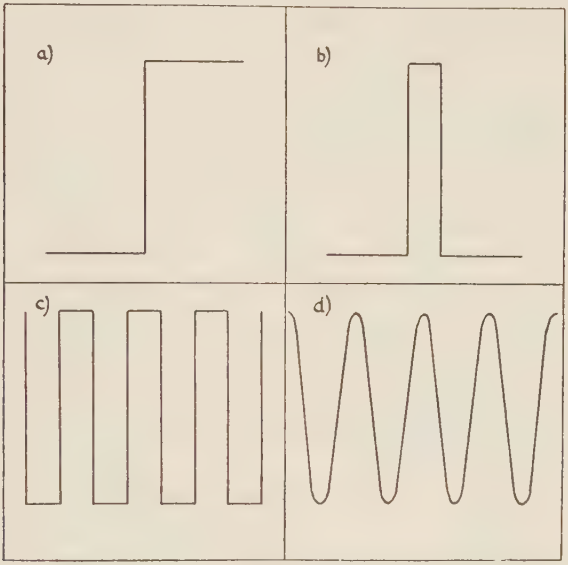


Figure 11

Functions for the distribution of illumination power in the ground-glass plane for the determination of areal faithfulness of photographic layers. (a) Step function (sharpness of contours). (b) Rectangular pulse (turbidity coefficient). (c) Rectangular curve (power of resolution). (d) Sinusoidal curve (frequency modulation).

from the resulting distribution of densities, with the aid of the density curve; analysing the effective distribution of light according to Fourier would give a certain dependence of frequencies, i.e. every single frequency of the continuous spectrum would be reduced in its amplitude by a definite factor of reduction. Taking instead of the "step function" the "rectangular pulse" or the "rectangular curve", we should again receive factors of reduction in the described manner; these factors should be identical for any chosen frequency, which means independence of the chosen distribution of illumination power. All these four functions of distribution are applied in practice, as is shown in Table II:

TABLE II

Function for the illumination power	Field of application
"Step function"	Sharpness of contours
"Rectangular pulse"	Turbidity coefficient
"Rectangular curve"	Resolving power
"Sinusoidal curve"	Frequency response in sound film

Unfortunately, all examinations carried out to date show that it is impossible to state unambiguously the frequency dependence of the factor of reduction by using only a single function, in order to calculate in advance the photographic reproduction of other distribution functions of the illumination power. If this were feasible only one measurement would have to be made, i.e. the frequency response would enable the calculation of the sharpness of contours, the turbidity coefficient or the resolving power. Even by eliminating the Eberhard, Ross, Kostinsky and other effects, it is impossible to arrive at an absolutely clear result. The best results are obtained by starting with the magnitude of the turbidity coefficient which is a measure for the scattering of light in the photographic layer and which is defined as the rise (astrogamma) of the so-called diameter-curve (turbidity curve according to Goldberg). Here, the broadening of any element of the image is, after Scheiner, approximated by the equation $b=b_0+\log(L/L_0)$. The fact that different functions were set up and experimentally verified (Greenwich formula: $\sqrt{b}$ instead of b; Ross formula: $\sqrt{(b+c)}$ instead of b) clearly shows the problem to be very complex. The toe in turbidity curves found by Goldberg[13] is due to the λ-dependence of k, according to Wildt[14]. Moreover, this author found that in the case of short exposures the resulting diameter may even be smaller than the required one, so that a narrowing of the elements of the image occurs.

Another important effect, first pointed out by the author[16], comes into play here. The diffused light in the photographic layer cannot be effective infinitely; thus a narrow beam of rays falling vertically on the layer may only influence a circle of the radius r_s (radius of scattering). With even exposure of a circle whose radius is smaller than r_s, the effective exposure must necessarily be smaller than that of a considerably larger area, due to the lack of a certain amount of scattered light from the surrounding area. Therefore, the density must be a function of the size of the exposed area element. This can easily be calculated. Assuming for the lateral diminuition of the scattering of light in the emulsion a function of e, i.e. $\vartheta(\xi)=\exp(-a\xi)$, which conforms quite well to practical conditions, we obtain a distribution of light for an image area of a slit of the width s_0, which is within the image of the slit,

$$L_\xi = L_0 \left\{1-\exp(-as_0/2)\left[\exp(a\xi)+\exp(-a\xi)\right]/2\right\},$$

and outside the image of the slit

$$L_\xi = L_0\exp(-a\xi)\left[\exp(as_0/2)-\exp(-as_0/2)\right]/2.$$

The distribution for different values of a is shown in Figure 12. It is seen that the intensity in the centre of the image of the slit diminishes with diminishing a. Substituting $\xi=0$, we obtain

$$L_{\xi=0} = L_0\left[1-\exp(-as_0/2)\right],$$

and find that the intensity in the centre is a function of the width of the image of the slit and the coefficient of scattering a. This dependence is illustrated in Figure 13.

From the assumption of an exponential diminishing of the scattered light follows also the formula for the turbidity curve

$$b = b_0 + (2/\alpha \log e)\log(L/L_0) = b_0 + k\log(L/L_0).$$

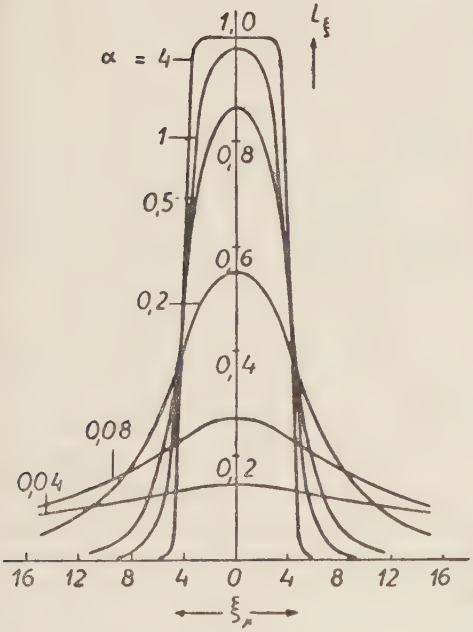


Figure 12

Curves for the distribution of light effective in the photographic layer by exposing an area, of the width s_0, extended to infinity in one direction. Assumption of a symmetrical exponential decline of light according to $\delta(\xi)=\exp(-\alpha\xi)$. The different curves are for different values of α which is a constant of the emulsion.

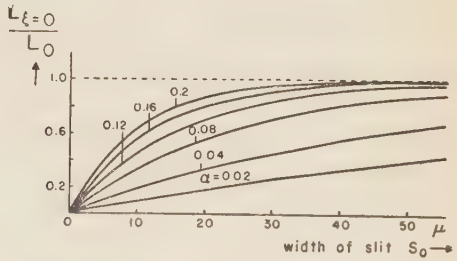


Figure 13

Dependence of the effective light in the centre of an evenly exposed area, of the width s_0, extended to infinity in one direction; for different values of the emulsion constant α.

With diminishing size of the exposed area elements we will have to reckon with diminishing density. From this follows the consequence, important in practice: *With an object to be photographed, containing area elements of varying sizes, we shall obtain a number of density curves, each of which pertains to one area element of particular size, and which are not parallel in the ideal case.* It may of course happen that at a given exposure the densities of large area elements fall into the straight part of the density curve, and those of small area elements lie in the toe or even below.

Another consequence arising from the turbidity curve concerns the problem of the effect of widening of an area on printing the negative. A widening in the negative may be compensated by an identical widening in the positive, a procedure which gained great importance in sound film production in the elimination of the so-called "thunder-effect".

From the above formula for the turbidity curve there follows for the widening $b-b_0 = k\log(L/L_0)$. Substituting $D = \gamma\log(L/L')$ and $D_0 = \gamma\log(L_0/L')$ ($\log L'$ = value of the inertia), then

$$\Delta b = b-b_0 = (k/\gamma)\,(D-D_0).$$

If this widening of an area element in the negative be eliminated by an equal widening in the positive, then

$$k_N(D_N - D_{0N})/\gamma_N = k_P(D_P - D_{0P})/\gamma_P;$$

this condition has been deduced by the author and has proved its worth in practice. We have, however, to pay attention to the simplifying assumptions in the deductions. k being a constant of the emulsion, D depends on the exposure and γ on the kind of emulsion and development, it can be quickly found in what manner a certain widening can be eliminated on printing.

In pictorial photography, always $\gamma_N < \gamma_P$ and $k_N > k_P$; then, D_P should be greater than D_N. Taking for the copy the same material as for the negative and $\gamma_N = \gamma_P$, we should assume $D_P \approx D_N$. Although these are obviously only approximations, they have been found very useful in sound photography. In any case, *a true transfer of areas is possible only by keeping to the rule that a widening of the area elements in the negative is compensated by an equal widening in the positive.*

The problem of sharpness of contours, i.e. the visibility of abrupt changes in density on the borders of area elements, is of special importance. Apart from the subjective aspect of the problem we have to reckon with two physical factors in the reproduction of an object: the first is the refraction and the second the aperture of the objective, which itself is assumed to be ideal, having no aberration. The refraction always causes a rounding at the edges of abrupt changes of light and can, for practical purposes, often be neglected. On the other hand, the rôle of the aperture of the objective becomes the more important, the thicker the emulsion is. This influence was thoroughly investigated by this author and Schimmel[12]. Turbulence then causes an additional bluntness in the photographic emulsion, which in turn leads to dullness. The maximum rise σ_{max} of the density curve in this transitory region was named "factor of sharpness" by Goldberg[13]; then

$$\sigma_{max} = dD/dx.$$

Substituting $D = \gamma \log(E/E_0)$ and remembering $k = db/d\log(E/E_0)$, then

$$\sigma_{max} = \gamma/k.$$

This equation, as set up by Goldberg, easily explains the fact that a Lippmann-plate, in spite of its small γ-value, gives a good reproduction of sharpness, due to little scattering. This relation of course only applies within the framework of the simplified assumptions made. Fine grain layers were recently investigated by Frieser[18]; the introduction of two scattering functions with altogether three constants (k_1, k_2 and ϱ as constant of the weight in the second function) resulted in a close approximation to the experimental diagnosis, without changing Goldberg's fundamental reasoning. We come now to the important problem as to how the eye perceives abrupt changes in density, or, in other words, whether we can give a physically measurable value for the subjectively perceived sharpness.

This last problem can be solved only by having abrupt changes in density of varying manner and form evaluated by as many viewers as possible, and by trying to interpret the results theoretically. Such investigations were carried out by Wolfe and Eisen[19]. They photographed knife-edges on 10 different negatives and then made contact prints on fine-grained positive material. Photometric measurements resulted in the curves shown in Figure 14. The images (diapositives) were evaluated by a

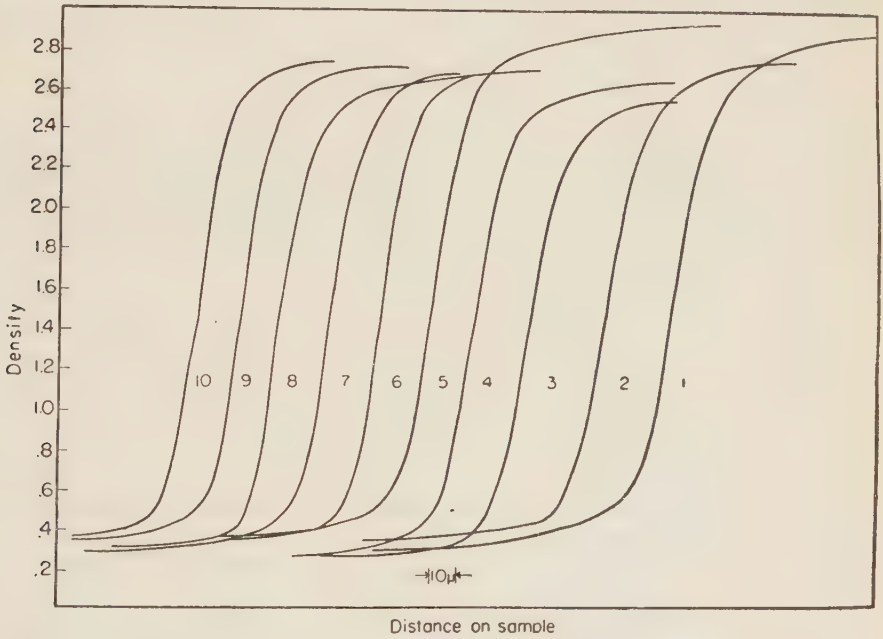


Figure 14

Densities on the edge of an exposed area, according to Higgins and Jones. The 10 curves correspond to 10 different negatives, all of which were printed on the same positive material.

number of viewers who sorted them in the order shown from 1 to 10, from smallest to greatest sharpness. It is seen that the maximum gradient does not vary greatly and it can be assumed that the factor of sharpness $\sigma_{max} = dD/dx$ is hardly a satisfactory measure of the perceived sharpness. The curves differ from each other by the form of toe and shoulder and also somewhat by the magnitude of the maximum difference of density. Higgins and Jones[19] used these experimental results in trying to find a physically measurable value satisfying the subjective findings. They found that neither the maximum gradient $(dD/dx)_{max}$ nor the medium gradient $(dD/dx)_{med}$ conform to the experimental results. Starting with thorough investigations of the graininess of photographic emulsions, they developed a theory according to which the physiological nystagmus (the irregular oscillation of the eyeball) causes the image projected by the eye-lens to wander in jerks over the retina and excite the sensitive elements of the retina according to the temporal changes in luminance dB/dt. The expression

dB/dt is the product of way-time change dx/dt multiplied by the simultaneously occurring change in luminance dB/dx.

Owing to the Weber-Fechner law, dB/dx has to be replaced by $(d\log B)/dx$, which in turn corresponds to dD/dx, the latter being the relative gradient at the edge of the density transitions. The subjective threshold value of the gradient dD/dx, as found in investigations on graininess, is 0.005 μ. As neither the mean nor the maximum gradient properly reproduced the observations, the following mean of the squares was formulated:

$$(G_x^2)_{Av} = \int_A^B (dD/dx)^2\, dx/(x_b - x_a).$$

The limits A and B are given by the threshold value of the gradient $dD/dx = 0.005\ \mu$. Furthermore, the subjectively perceived sharpness also depends on the difference of densities ΔD; thus the value $Ac = (G_x^2)_{Av}$, ΔD was formulated and termed "acutance". This value conforms well with the visual evaluation, as seen in Figure 15, in which

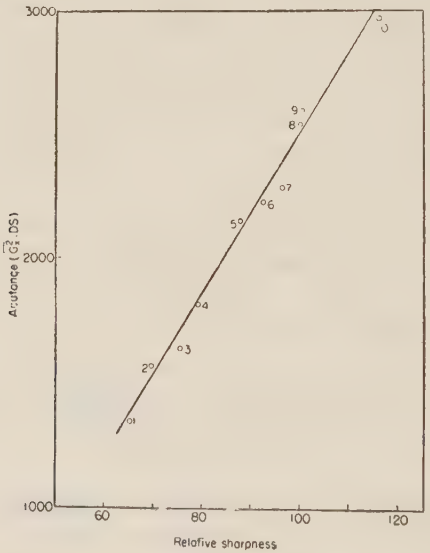


Figure 15

Comparison of the subjectively evaluated knife-edge tests of Figure 14, with the objectively determined magnitude of the sharpness of contours (acutance).

the physically measured values are superimposed on the subjectively found relative values of sharpness. These results show that *the chosen mean value of the gradient strongly depends on the special form of the density at the edge, and the bends of the curve apparently play a bigger role than the straight middle part.*

The density ranges from fog to a very high value; therefore the toe and the shoulder of the density curve were certainly used and they were certainly different for each of the negatives. The important question now arises as

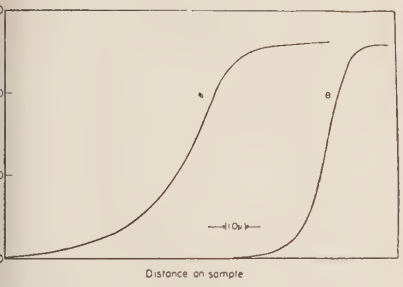
to whether the 10 negatives differ in their functions of scattering to such a degree, that they explain the findings illustrated in Figure 15, or whether these findings are a result of the distribution of the density at the edge, which depends on the form of the density curves of the negatives.

As a first approximation, it was possible to characterize the scattering properties of an emulsion by the turbidity coefficient k. The question should be asked, therefore, whether the emulsions of the 10 negatives also possess 10 different values of k and whether k may be substituted for Ac in order to obtain close conformity with the visual perception of sharpness. This was recently proposed by Frieser[18]. Higgins and Jones regrettably did not deal with this question, but the form of the density curve is probably of decisive importance, especially the toe and the shoulder, as the gradients of the straight parts of the transitions of densities in the 10 examined cases do not differ greatly one from the other. It should thus be concluded that the value of k alone is insufficient for obtaining a measure of the visual perception of sharpness, and the density curve has also to be included in the considerations. Although the product of the gradients $\gamma_N \cdot \gamma_P = 1$ in the negative-positive process, it was proved in the preceding section that a greater value is preferred in general if a "first choice" copy is wanted, with express use of the toe and shoulder of the density curve of the positive. The exploitation of bent parts of the curve has also to be expected in the negative. *Therefore, the entire process of transfer has to be included in the considerations; thus, a statement on the reproduction of details of sharpness and its importance for the subjectively perceived sharpness can only be made by simultaneously considering the reproduction of details of brightness, which in turn depends on the form of the copy density curve.* The significance of these considerations can be clearly recognized by looking at the conclusions drawn by Higgins and Jones from their investigations and which deal with the problem of conformity of sharpness and resolving power. They printed a negative on two positive emulsions of different resolving power, i.e. 230 and 130 lines per mm, in such a manner as to obtain the best-matched tone reproduction possible. Figure 16 shows the most characteristic parts of the published images (three-quarters of the originals).

Higgins and Jones concluded from these images that the positive with the higher resolving power possesses an unsatisfactory sharpness of contours, whilst the opposite applies to the positive with the lower resolving power. Knife-edge images of both positive emulsions are also reproduced in Figure 16 and it is obvious that curve B will render a greater sharpness of contours than curve A. However, the arguments of Higgins and Jones cannot be accepted. Viewing the two images, one finds that Figure 16a is far from being a "first choice" copy. It is contrasty to such a degree that all parts of the wickerwork in the near lower part of the chair have completely disappeared. The shadows exclude all details, and so do the lights. The images cannot be compared at all with each other, because an even slightly satisfying reproduction of the object range was apparently impossible in Figure 16a. *Two emulsions can only be compared with regard to their ability of reproducing details*

of sharpness by creating conditions conducive to a satisfactory reproduction of details of brightness. This principle has also to be preserved when defining the resolving power. It is useless to quote the maximum resolving power when it is not near the densities which pertain to the practical conditions of reproduction. No statement regarding the quality of reproduction is possible without knowledge of the complete characteristic line of reproduction. Unfortunately, Higgins and Jones did not quote the γ-values of the two experimental emulsions and the copy density curves which served as basis for the two images. Moreover, it would be interesting to know the values of the resolving power which would have been obtained for the 10 negatives of Figure 14 which were copied on the same positive material. It may be safely said that curve 10, which gave the greatest sharpness of contours, would also have rendered the highest resolving power. It cannot be understood why the emulsion possessing such a great sharpness of contours (Figure 16b, curve B) should have less resolving power than the emulsion with the entirely dull transition at the contours (Figure 16b, curve A). A large number of experiments is still needed to clarify this problem finally. In these experiments characteristic lines of reproduction and practical photographs should be made in addition to knife-edge and resolving tests. They should then be evaluated and visually judged in order to pick out the best copies, as was done by Goldberg and later by Higgins and Jones and many others. The simultaneous measurements of sharpness and resolving tests will then serve as a basis for stating the conditions for optimum reproduction of details of sharpness and brightness. Both these characteristic values of the photographic image are interdependent, because the degrees of brightness in an object will almost always be accompanied by the occurrence of contours. It is therefore necessary to treat both problems at possibly the same time and to investigate their interaction.

In this treatment we have to consider the already fully discussed points of view. When trying to solve simultaneously the problems of reproduction of details of brightness and sharpness, one may be tempted to proceed in analogy to television technique. The entire image area to be reproduced is subdivided into individual lines which have to be imagined to be joined one to the other; then we inquire into the conditions for the reproduction true to nature of the distribution of brightness by adapting the height of the lines to the highest detail of sharpness still to be reproduced. To the distribution of brightness impressed onto the photographic emulsion corresponds a distribution of densities depending on the sensitivity of the emulsion and its other properties; the mean density corresponding to the mean illumination power is modulated differently according to the object to be photographed. The task therefore consists of reproducing the mean value and the modulation, for which frequency, amplitude and phase are decisive. Here enters the demand to keep the noise level sufficiently small, which, in photography, equals the graininess of the emulsion. By disregarding this factor at first, the problem can be completely solved in the case of electric reproductions, when the characteristic line of reproduction is known in addition to the frequency- and phase-dependency of the whole



(b)

(a)

(c)

Figure 16

Comparison of the qualities of reproduction of a negative printed on two different positive materials. (a) Positive material of high resolving power (230 lines per mm). (b) Knife-edge tests of the two positive materials (left: 230 lines per mm; right: 130 lines per mm). (c) Positive material of lower resolving power (130 lines per mm).

system. The manner of reproduction of any given distribution can then be stated in advance. The conditions applying to photography are slightly different. There is a definite facilitation in the fact that in principle no shifting of phases is to be expected. On the other hand, matters are made more difficult as a series of characteristic curves has to be reckoned with. The reasons for this plurality are the following:

(i) The characteristic curve depends on the frequency because the effective illumination power is a function of the size of the exposed area element. In trying to conclude from any given distribution of densities the distribution of the effective illumination power, the basic curve has to be established from a gradation of densities whose area elements are of a size equal to that of the distribution of densities to be examined. It must not be taken for granted that the density curves corresponding to differently sized areas are parallel.

(ii) The characteristic curve depends on the frequency in this way: owing to the scattering of light in the layer, the amplitude diminishes with increasing frequency causing an effect of rectification, i.e. a rise of the mean density. As opposed to electric characteristic curves, this effect of rectification may originate without distortion and a superimposed sinusoidal oscillation of light may remain undistorted. By appropriate selection of the conditions of printing, one succeeds in compensating the rectification effect for a given frequency, but not for all frequencies simultaneously. Moreover, the rectification effect is a function of the modulation, as it increases with rising modulation. It therefore follows that in pictorial photography it is impossible to secure a simultaneous reproduction true to nature of all details of sharpness. *The ideal case of a faithful transfer of area cannot be achieved in principle because the image consists of a great number of area elements of varying size of widely differing distribution.*

The following consequences result from the foregoing considerations: The eye is the highest authority in judging the quality of an image. When a physically correct reproduction can only be approximated, we have to find the permissible and not disagreeably perceived deviations. Just as Goldberg thoroughly investigated the problem of the importance of details of brightness and solved it with the aid of numerous practical experiments, one has to inquire into the importance of details of sharpness. Investigations on these lines, however, have not yet been carried out. Apart from the problem of reproducing line drawings, the problem of reproduction of details of sharpness of half-tone patterns cannot be separated from that of details of brightness. Thus, only copies can be evaluated due to the demand for the correct reproduction of all the brightnesses in the object. Although Higgins and Jones recently proved in their basic investigations that even by deviating from the conditions to be imposed from a physical point of view, "first choice" copies may be produced, it is found, on the other hand, that the entire process of transfer from negative to positive has to be approximated, as far as possible, to the ideal conditions. It is therefore indispensable to find the characteristic curve of transfer of the negative-positive process

and to adapt this curve to the object range to be reproduced. The same applies to the transfer of details of sharpness. Here, too, considerations have to start with a given object range which has to be transferred as correctly as possible in respect of its brightnesses. Only then can the manner of reproduction of details of sharpness be stated. This question is important because the reproduction of details of sharpness with inclusion of the copy has been thoroughly examined only in sound photography, while no such investigations have so far been made for pictorial photography. In the latter, only different emulsions are compared as to sharpness or resolving power without considering whether conditions conducive to the production of a satisfactory copy were chosen. The disregarding of this point of view is certainly the reason for the conclusion drawn from the investigations of Higgins and Jones that sharpness of contours and resolving power do not run parallel. The resolving power is — due to the determination of a threshold value — certainly no ideal value of characterization; it would be advantageous to give the frequency response as this would afford a statement on those line spacings which do not conform with the resolving limit. It has been found, however, that with the aid of the Schumann test[12] one arrives at rather reproducible values; in addition, this author showed in his investigations that with a high resolving power, and therefore at a high frequency limit, there is also a greater amplitude at lower frequencies, which means a better frequency response. With a better response, we can also assume a greater sharpness of contours. As the turbidity coefficient k in the first approximation defines the frequency response and thus the resolving power, it cannot be understood why sharpness of contours and resolving power should give partly controversial results. Higgins and Jones proved that in the case of acutance a mean of the squares of the gradients for those points of the curve which lie in the transition zone of a knife-edge test, conforms well with the visual diagnosis. They were the first to evolve the experimental foundation of the subjective side of the problem and to express it by means of a formula. But the improbability has already been pointed out that in the ten considered cases the turbidity coefficient k has changed to the same extent as the value of Ac defined by them, because the latter is arrived at on the basis of physiological viewing conditions, whilst k refers to the objective finding of the influence of scattering. k could be a measure of the sharpness of contours if it would change in the same sense as Ac. This appears improbable according to the investigations made by Higgins and Jones, as the subjective impressions are apparently greatly influenced by the form of the curve in the transition zone, and this form is not covered by the constant k. It is to be remembered, moreover, that k is constant only within a certain range; the diameter curve, always bent in its lower part, is bent in its upper region also and is supposed to approach a limit, according to recent investigations made by Vigon and Olagorta[20].

REFERENCES

1. Goldberg, E., 1922, *Der Aufbau des photographischen Bildes*, Knapp, Halle (2nd ed. 1925).
2. Narath, A., 1933, *Kinotechnik*, **15**, 341—344, 363—366, 382—384, 401—404.
3. Porter, A. W. and Slade, R. E., 1919, *Phil. Mag.*, **38**, 187—197.
4. Hurter, F. and Driffield, V. C., 1890, *J. Soc. chem. Ind., Lond.*, **9**, 455—469; 1891, *ibid.* **10**, 100—104.
5. Jones, L. A., 1920, *J. Franklin Inst.*, **190**, 39 ff.; Mees, C. E. K., 1942, *The Theory of the Photographic Process*, New York.
6. Jones, L. A., 1949, *Photogr. J.*, **89B**, 126—151.
7. Lichte, H. and Narath, A., 1941, *Physik und Technik des Tonfilms*, Hirzel, Leipzig (3rd ed. 1945).
8. Pistor, W., 1934, *Z. tech. Phys.*, **15**, 107—112.
9. Narath, A., 1935, *Kinotechnik*, **16**, 51—55.
10. Cherry, E. C. and Gouriet, C. G., 1953, *Proc. Instn elect. Engrs*, (100), 9 ff.
11. Frieser, H., in Stenger-Staude, *Fortschritte der Photographie*, Leipzig, 1940, **2**, 277—366; 1944, **3**, 73—112.
12. Narath, A. and Schimmel, G., 1953, *Proc. Int. Congr. Photogr.* Paris, 1951, 280—294; 1953, *Proc. Int. Congr. Photogr.* London, 1953; 1955, *Optik*, **12**, 205—214; 1955, *Sci. Industr. Photogr.*, **26**, 345—353; 1952, *Phototech. u. -wirtsch.*, (11), 439—440; *ibid.* (12), 482—483.
13. Goldberg, E., 1913, *Z. Photogr.*, **12**, 77 ff.
14. Wildt, R., 1928, *Z. Photogr.*, **25**, 153 ff.
15. Narath, A., 1938, *Kinotechnik*, **20**, 93—96.
16. *Idem*, 1934, *ibid.* **16**, 351—354.
17. *Idem*, 1937, *ibid.* **19**, 305—310.
18. Frieser, H., 1955, *Mitt. ForschLab. Agfa, Leverkusen-Muenchen*, 129—152.
19. Higgins, C. C. and Jones, L. A., 1952, *J. Soc. Mot. Pict. Tel. Engrs*, **58**, 277—290.
20. Vigon, M. T. and Olagorta, M. A., 1955, *An. Soc. esp. Fis. Quim.*, **51**, 193—214; ref. 1956, *Sci. Industr. photogr.*, **27**, 102—104.

ON SOME PROPERTIES OF PHOTOGRAPHIC LAYERS FOR THE PRODUCTION OF MIKRATS

H. Frieser

Wiss. Photogr. Labor. der Agfa Aktiengesellschaft für Photofabrikation, Leverkusen-Bayerwerk

ABSTRACT

The method proposed by Goldberg for the production of extremely small photographic reductions ("mikrats") is discussed. By using emulsions of silver bromide with development, times of exposure could be shortened. Some properties of these emulsions are discussed.

1. At the International Congress for Photography held in Paris in 1925, Goldberg demonstrated photographs of a size then considered unattainably diminutive[1]. All the details of a portrait of Nieps, having a diameter of approximately 0.1 mm, could be discerned under a highly magnifying microscope. Considerable photographic reductions as such had been known for some time. These had gained some importance thanks to Stanhope who had used them in delivering dispatches from besieged Paris during 1871. The scale of reduction of these pictures was surpassed by a factor of 10 by the reductions produced by Goldberg, who called his innovation "mikrats", a name which is gaining more and more popularity.

Goldberg saw the major importance of his method in its application as a tool for librarians by helping to control the rising tide of printed matter. For him the library of the future consisted not of archives filled with books, but of stores of small boxes of mikrat plates, each containing several books which could be read by means of special viewers. It would, of course, be difficult to find a desired work from among a large number of documents. Goldberg therefore devoted much thought to photoelectric methods of registering photocopies — a problem which has recently again received much interest — and proposed an appropriate procedure[7].

In order to obtain good mikrats, a series of optical and photographic difficulties had to be overcome. The first, consisting primarily in the production of suitable photographic objectives, will be dealt with only slightly. Goldberg used microscope objectives, but these gave sharp definitions within a small angle only, so that the reproduction of large-sized newspapers or drawings sometimes presented obstacles.

The photographic difficulties consist in the reproduction of a layer working practically grainlessly and having a resolving power sufficient to reproduce the minute details of a mikrat.

Received June 27, 1956.

To achieve this, Goldberg used silver chloride emulsions which blacken directly without development. After toning with gold, very sharp images resulted. These layers had the disadvantage of being very insensitive, which is understandable with non-development layers. Therefore, as described previously[2], I tried to produce silver bromide layers which had to be developed and also possessed sufficient resolving power for mikrats. With these layers it became possible to reduce a magazine page from 21 × 30 cm to ~ 1 mm². 10,000 of these images could thus be placed on an area of 9 × 12 cm. A comparison of the volumes of original and reproduction is interesting. Taking the average thickness of paper as 0.1 mm, the volume of 10,000 pages will be 60,000 cm³, the volume of the reductions (plate thickness 1 mm) being 10 cm³. Thus, the ratio is 1:6,000.

Although the sensitivity of this layer is as yet low, it became possible to produce mikrats with exposures of 1—2 seconds, with incandescent light and focal apertures not exaggerated (approx. 2.5). A negative on 35 mm film served as starting point.

Similar emulsions are currently produced by Agfa ("Mikratplatte") and Kodak ("Maximum Resolution Plate").

2. As mentioned above, an emulsion suitable for the production of mikrats has to have very fine grain and also good resolving power. A mikrat having 500 lines per mm requires good resolution of lines at intervals of 2 μ. This is approximately four times the wavelength of light, and shows the extreme demands made on the objective. The demands on the photographic emulsion are equally severe.

The resolving power of a photographic emulsion is primarily limited by haloes due to diffusion. The diffusion haloes cause an exposure made on the surface of the emulsion to affect also adjacent areas not receiving direct light. This is due to the scattering of light by the silver bromide grains. Therefore the separation of exposed and unexposed areas is vague and not sharp, and the contrast is further reduced when a line screen is reproduced. When the lines of the screen are very close, the contrast becomes so small as to defy resolution of the screen. The limit of resolution reached gives the resolving power of the emulsion. For normal layers used in pictorial photography, this will be about 100 lines per mm or intervals of 10 μ between lines, which is much higher than the value required for mikrats. Although we postulated the above intervals of 2 μ for sufficiently good reproduction, power of resolution is defined by a reproduction which under good viewing conditions just allows for resolution. In order to reproduce a screen with intervals of 2 μ, a power of resolution substantially higher than 500 lines per mm is a precondition.

This aim can be achieved by reducing diffusion haloes, which means using layers having so fine a grain that light is not appreciably scattered. It was found that this is attained only with a mean grain diameter equalling half the wavelength of light, i.e. 0.2—0.3 μ. Larger grains cause turbidity and diffusion haloes.

An efficient method to characterize the diffusion halo[3,4] was also used by Goldberg[5]. Strips of identical widths are photographed with increasing exposure times. It is

obvious that the diffusion halo will be large when the width of the image exceeds that of the strip and it rapidly rises with increasing exposure times. It was found that the width of the image increases linearly with the logarithm of the exposure H[3,4]:

$$b = A + k_2 \log H.$$

Normal layers behave as shown in Figure 1. There the width of the strip is superimposed on the logarithm of the exposure. It is seen that after a short slight rise the width of the image of the strip increases linearly.

Glass plates were coated either immediately or after various lengths of ripening at elevated tempreatures with emulsions which had been precipitated so quickly that they remained transparent and caused no diffusion. The unripened plate had a grain diameter of about 0.02 μ and was almost perfectly transparent. However, the longer the period of ripening, the greater became the grain diameter and turbidity. The strips were then exposed onto the plates and the widths of the images were measured following development. Results are shown in Figure 2.

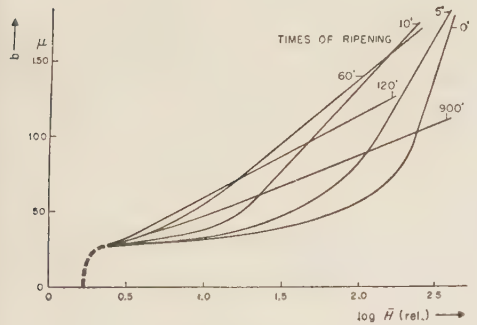


Figure 1

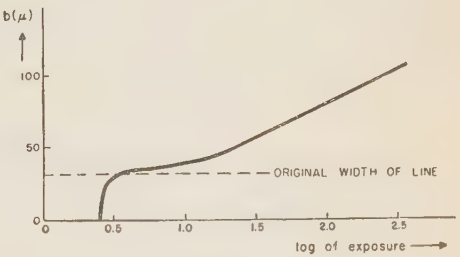


Figure 2

It can be seen that for unripened emulsions, the width of the image of the strip remains constant over an extended range ($\Delta \log H = 1.5$!), but it increases rapidly beyond this value. As this rise applies only to short exposures, it is normally not met with. Thus, the high resolving power of the emulsion, which reaches about 1,000—1,300 lines per mm, is explained. These figures refer to values measured with microscope objectives. With interference methods similar layers showed resolving powers of 2,500 lines per mm.[8] With increased ripening, the range in which the width of the image remains constant decreases, and the rise becomes flatter. It now overlaps normally used exposures, and resolving power decreases.

With still further increased times of ripening, the range of unchanged image widths disappears almost completely, the rise is slow and corresponds rather to a normal emulsion (Figure 1).

These curves demonstrate that high resolution is possible only with perfectly clear emulsions. Even a short ripening, to improve sensitivity, will cause turbidity and always reduce resolution.

The facts presented in Figure 2 can be interpreted as follows: Diffusion in unripened emulsions is small. However, as soon as it exceeds the threshold, due to long exposure, its effect will be rather far-reaching, as the clear and non-diffusing emulsion absorbs only little diffused light. The image width then increases rapidly with the logarithm of the increase of exposure. With longer ripening, turbidity and the amount of diffused light increase. The effect of the latter will already be felt with short exposures. But the diffused light will be more rapidly absorbed by the emulsion, so that the straight part of the curve becomes flatter. The more the ripening has progressed, the more will this phenomenon be marked.

3. So far we only discussed diffusion haloes. Now, we shall also add some remarks on sensitivity.

It is obvious that the sensitivity of mikrat layers is extremely small. This is a basic fact. Although sensitivity can be raised somewhat by means of suitable sensitizers, it is limited by the minute quantity of light a small grain is able to absorb.

Experiments carried out in our laboratory[6] established that sensitivity increases approximately with the cubes of the grain diameters, i.e. the volume of the grains, assuming them to be spherical. The theoretical relation of sensitivity for a layer of grains of 0.8 μ will therefore be 1:64,000.

REFERENCES

1. Goldberg, E., 1925, *C.R. VI Int. Congr. Photogr.*, Paris; 1926, *Photogr. Ind., Berl.*, 484; 1926, *Z. tech. Phys.*, **7**, 500.
2. Frieser, H., 1941, *Z. angew. Photogr.*, **3**, 57.
3. *Idem*, 1955, *Photogr. Korr.*, **91**, 69; 1956, *ibid.* **92**, 51.
4. *Idem*, 1955, *Mitt. Forsch. Lab. Agfa Leverkusen-Muenchen*, **1**, 129.
5. Goldberg, E., 1912, *Photogr. J.*, **36**, 300.
6. Frieser, H. and Klein, E., 1954, *Z. Elektrochem.*, **58**, 655.
7. Goldberg, E., 1931, *Ber. VII Int. Congr. Photogr.*, Dresden, 317.
8. Wolfe, R. N. and Eisen, S. C., 1950, *J. opt. Soc. Amer.*, **40**, 143.

THE CYANINE SENSITIZING DYES (A REVIEW)

C. E. Kenneth Mees

Honolulu, Hawaii

Photographic materials are sensitive only to the light that they absorb. Silver chloride, which is almost colourless, is sensitive only to ultraviolet light, which it absorbs. Silver bromide, which is pale yellow, is sensitive to violet and blue light. Thus early photographic materials were sensitive only to light of the shorter wavelengths — ultraviolet, violet, and blue. The discovery that sensitiveness to other parts of the spectrum could be produced by the addition of dyes was made in 1873 by H. Vogel.

Vogel's discovery excited considerable attention, and many experimenters attempted to repeat his results. Some were successful; others, unsuccessful. It gradually became known that to obtain a good sensitizing effect, the dye must be very pure and used in a dilute solution, and at that date the structure of dyes and their chemical preparation in a pure state was little understood. Moreover, all dyes do not possess sensitizing action, and many give rise to side effects which spoil the emulsion. The early work was done on collodion emulsions, and it was not until 1882 that commercial dry plates were sensitized, with the dye eosin.

In 1883 J. M. Eder investigated systematically dyes of the eosin series and found that erythrosin (tetraiodofluorescein) was a valuable sensitizer for the green region of the spectrum with gelatin emulsions and was superior to eosin. Erythrosin was adopted for the manufacture of so-called *isochromatic* or *orthochromatic* plates and continued in use almost to the present time.

In 1884 Vogel found that a mixture of quinoline red and cyanine would sensitize throughout the spectrum as far as the orange. A great deal of work was done on sensitizing by dyes, but no advance of importance was made until 1901, when Miethe and Traube prepared dyes similar in chemical structure and in methods of preparation to the cyanine used by Vogel, and which were greatly superior to cyanine as sensitizers for the yellow and orange, particularly because they had less deleterious effects upon the photographic properties of the emulsion. One of these dyes which was particularly valuable was called *ethyl red* and represented a notable advance in sensitizing dyes. At almost the same time as the work of Miethe and Traube, Ernst Koenig, who was a chemist at the Hoechst Dye Works, made a number of dyes of the same type as ethyl red and placed them on the market as *orthochrome T*, *pinaverdol*, and *pinachrome*. All of these dyes were made by the condensation in the presence of alkali

Received June 10, 1956.

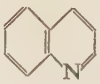
of quinoline salts with quinaldine salts, the quinoline or quinaldine nuclei containing substituents which determined the exact structure and properties of the dye. Thus pinachrome contains two ethoxy groups. Up to that time, no entirely satisfactory sensitizer for the red was known, but in 1905 B. Homolka, one of Koenig's colleagues in the dye works, found that a very effective red sensitizer could be prepared by carrying out the condensation in the presence of formaldehyde or chloroform. The new product, named *pinacyanol*, was recognized as belonging to a new and distinct group, although its structure remained obscure. Pinacyanol was the dye used in making the first bathed plates at Wratten and Wainwright in Croydon, England, in 1906. The Wratten and Wainwright Panchromatic plates were sensitized with a combination of pinacyanol and pinachrome.

During the First World War, the impossibility of importing German dyes turned the attention of British and American chemists to their synthesis. In England, research on sensitizing dyes was carried out in the laboratories of organic chemistry at Cambridge University by W. J. Pope and W. H. Mills. As a result of this work, the structures of the cyanine dyes then known were firmly established, and notably the difference in structure between the cyanines and the carbocyanines was determined. Mills, moreover, succeeded in making cyanine dyes from other bases than those derived from quinoline. Thus he prepared an analogue of pinacyanol from the base benzothiazole, which contains sulphur, and found that it had a powerful sensitizing effect with less tendency to fog the emulsion than pinacyanol.

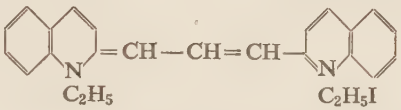
In 1923 the English firm of Ilford, Ltd., decided to add a dye chemist to their staff, and Sir William Pope suggested F. M. Hamer, who had been working with Mills in the cyanine dye field. In the Ilford Laboratories, Dr. Hamer not only did research work, but she published a number of important papers. One of these, published in the Journal of the Chemical Society, dealt with an improved synthesis of carbocyanines and described a number of carbocyanines prepared by the new method. Another paper was published in the Photographic Journal for January 1928, and in this most interesting paper she gave the absorption and sensitizing curves for one from each of the known classes of cyanines, fifteen dyes in all. Organic chemists in England and the United States realized that by varying the bases and the intermediate products used in synthesis, a great number of sensitizing dyes of the cyanine type could be produced.

Cyanine dyes consist of two nuclei joined by a chain of CH groups. Since the carbon atom has a valency of four, a CH group has three free valencies available for its attachment: $=CH-$. In dyes, these valencies are arranged in what are called *conjugated chains*, the valency bonds being alternately single and double: $=CH-CH=CH-$. The conjugated chains are associated with the absorption bands, and therefore the colour of the dyes, because the valencies can change their positions in the chain, $=CH-CH=CH- \longleftrightarrow -CH=CH-CH=$, and the *resonance* of the molecule produced in this way gives rise to the absorption of light. The cyanine dyes are derived from nuclei composed of rings of atoms, which include nitrogen,

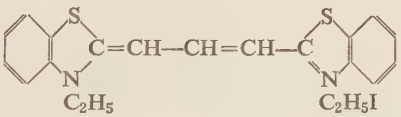
and the conjugated chain of single and double valency links starts and ends with the nitrogen atoms. Thus the first known cyanines were made with nuclei derived from quinoline:



Pinacyanol iodide was shown by Mills and Hamer to have the structure:

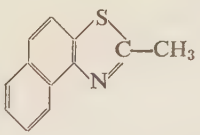


If dye is made by the same process as that used for making pinacyanol but, instead of quinoline, the base benzothiazole is used, a very clear, bright purplish red dye is obtained, the structure of which is:



This dye is an excellent sensitizer.

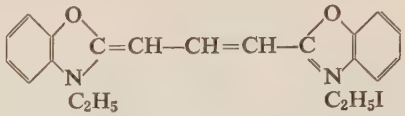
In 1928 L. G. S. Brooker in the Kodak Research Laboratories, Rochester, N. Y., prepared cyanine dyes derived from 2-methyl-β-naphthothiazole:



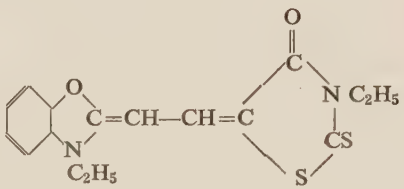
He obtained dyes which were excellent sensitizers and which absorbed red light and sensitized for the red. While experimenting with these red sensitizing dyes, the Kodak Research Laboratories found that a combination of some of the dyes derived from benzothiazole with one of the dyes made from naphthothiazole gave a very much stronger sensitizing than could be obtained by the use of the single dyes and even more than the total of the sensitizing power of the two separate dyes. Dye A, for instance, may be a red sensitizer; and Dye B, a green sensitizer. But it may be found that the use of Dyes A and B together may give a red sensitivity beyond that available by the use of Dye A alone.

In 1933 a new group of sensitizing dyes was discovered which differed from the cyanines in that part of the molecule contained an oxygen atom in place of the nitrogen of the cyanine nuclei. These new dyes were discovered at the same time by

Dr. Brooker and by J. D. Kendall in the Ilford laboratories in England and were named *merocyanines* at the suggestion of Dr. Hamer. Thus a typical cyanine (*oxacarbocyanine*) has the structure



whereas a typical merocyanine is



Merocyanines are produced by the condensation of an intermediate produced from a nucleus which could be used to make a cyanine with a compound containing a reactive $COCH_2$ group. A great number of compounds contain this group and, since each of them can be combined with intermediates from many nuclei, a vast number of merocyanines can be made. In practice, merocyanine dyes are useful especially as sensitizers for paper emulsions and as supersensitizers for green-sensitive emulsions. They include many valuable dyes, especially those which are solubilized by the inclusion of acid groups in their structure.

An important result of the discovery of sensitizing dyes is the extension of the spectral region which can be photographed. From a photographic point of view, the spectrum may be divided into four regions: the region from the ultraviolet to 5000 Å, which can be photographed on plates containing no dye sensitizers; the region from 5000 to 7000 Å, which includes the remainder of the visible spectrum and which can be photographed on panchromatic materials with short exposures; the region from 7000 to 9000 Å, the near infrared, which can be photographed on special materials with exposures greater, but not much greater, than those necessary for the visible spectrum; and the region beyond 9000 Å.

It will be remembered that in the cyanine dyes, two nuclei formed of rings of atoms and containing basic nitrogen atoms are joined to form a dye by a chain of methine (=CH—) groups. Heavier nuclei give dyes with absorptions and sensitizing maxima displaced towards longer wavelengths. Similarly, lengthening the chain of methine groups joining the nuclei moves the absorption toward longer wavelengths. The shortest chain consists of one methine group only, and the dyes are known simply as *cyanines*. The next chain has three methine groups, and dyes with this structure are termed *carbocyanines*. Dyes with five methine groups in the chain are known as *dicarbocyanines;* those with seven methine groups as *tricarbocyanines;*

with nine methine groups as *tetracarbocyanines;* and with eleven methine groups as *pentacarbocyanines* (Figure 1).

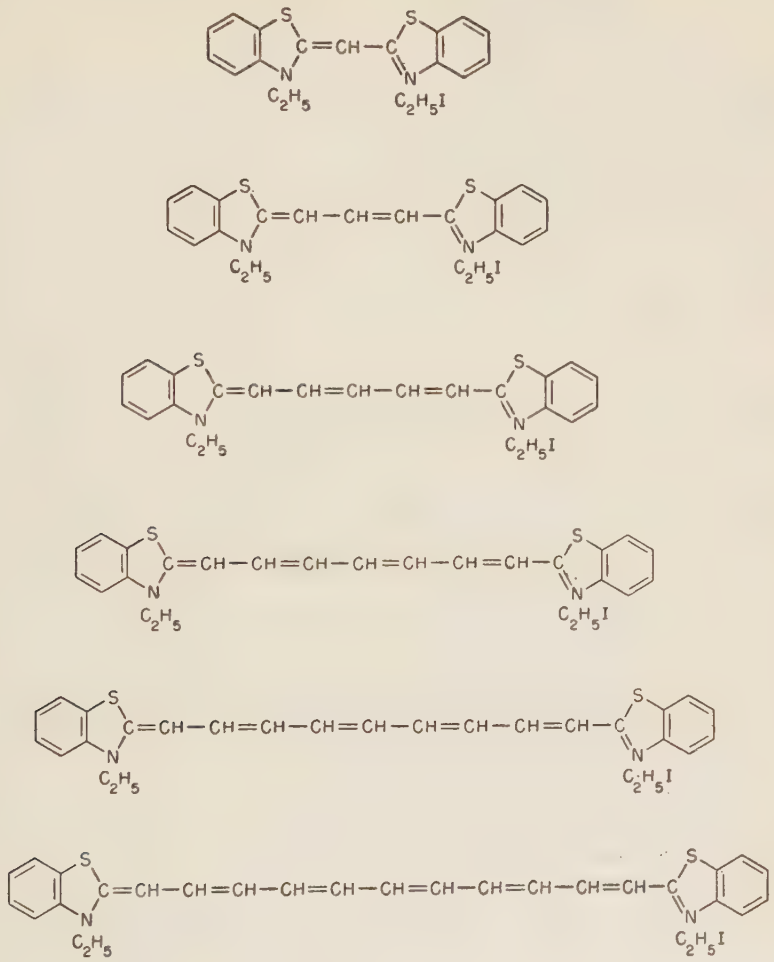


Figure 1

Cyanine dyes with chains of increasing lengths.

Figure 2 shows the progress which has been made in the extension of the spectral region for which photography can be employed in practice. At the top is the spectral region, including only the blue, violet, and ultraviolet, which could be photographed on silver-bromide plates without any sensitizer. Then the discovery of colour sensitizing by Vogel and particularly the use of erythrosin made it comparatively easy to photograph through the green region of the spectrum and record wavelengths up to approximately 6000 Å.

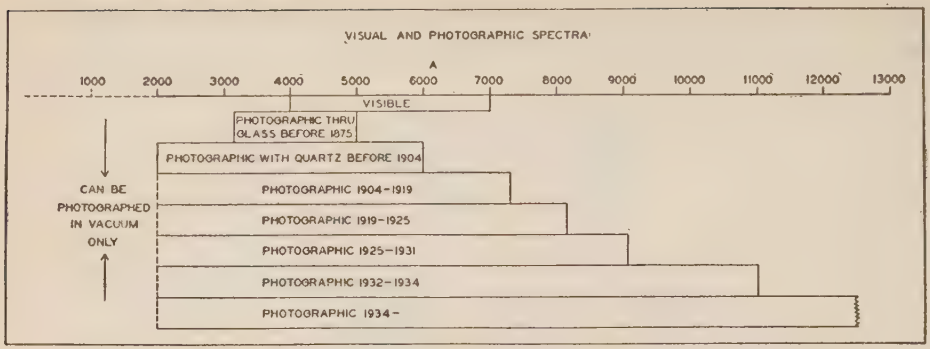


Figure 2

Chart showing progress in the photography of the spectrum.

In 1904 the application of Homolka's pinacyanol to the production of panchromatic plates made it possible to photograph to the limit of the visible red, a region which may be roughly placed as just beyond 7000 Å. In 1919 E. Q. Adams and H. L. Haller in Washington discovered the dye *kryptocyanine*, a carbocyanine from lepidine, with which sensitivity could be obtained up to beyond 8000 Å. In 1925 Dr. Hans T. Clarke found in the Kodak Research Laboratories that in some kryptocyanine preparations another dye, which was named *neocyanine*, was formed, by which the photographic spectrum was extended to 9000 Å. The discovery of the tricarbocyanines made it possible to make the tricarbocyanine from lepidine, and it was named *xenocyanine*. With the aid of this dye it became possible in 1932 to extend the limit of the photographic spectrum to 11,000 Å, using long exposures. In 1934 the tetracarbocyanines and pentacarbocyanines were made, with which the spectrum of the sun could be photographed to somewhat beyond 13,000 Å.

If, in the period between 1933 and 1940, an observer had made a survey of the available sensitizing dyes, they might have seemed to him to comprise a rich and heterogeneous assemblage, enormously more varied in type and capable of giving a far greater variety of sensitized photographic products than the dyes available at the beginning of 1928, but still a miscellany of dyes, nonetheless, individual members of which were in many respects quite unrelated to others. Anomalies of behaviour abounded, some having to do with the light absorption of the dyes, others with photographic sensitization, and others with the synthetical procedures used for obtaining the dyes.

As time went on, theories of the relationship between colour and chemical constitution, of photographic sensitization, and of the relationship between chemical structure and reactivity were gradually built up by the aid of which many of the apparent anomalies could be explained. These studies were doubtless started because they were irresistible to anyone working with the cyanine dyes possessed of a modicum of scientific curiosity, but experience was to show time and again that from quite disinterested speculation important practical developments very often arose.

The mechanism of sensitizing by dyes has attracted much attention, but only in recent years has any real advance been made in understanding it. The mechanism by which dyes sensitize silver halides to light of the wavelengths which they absorb may be summarized as follows:

Sensitizing dyes are adsorbed to the silver halide grains, the light absorbed by the dye molecules increases their energy, and this energy is transferred to the silver halide, where it appears in the form of free electrons, the final result being the formation of latent image, much as when the light is directly absorbed by the silver halide. The development of automatic photoelectric spectrophotometers made it possible to measure throughout the spectrum the light reflected and transmitted by a dyed emulsion and, therefore, to calculate the light absorbed. Using one of these instruments, J. A. Leermakers, B. H. Carroll and C. J. Staud in 1937 showed that the sensitizing spectrum corresponds precisely to the absorption spectrum of the dyed silver bromide.

The primary process in optical sensitization is the absorption of a quantum by a dye molecule with excitation of the dye molecule; the excitation energy is transferred to the silver halide, which liberates an electron from the lattice as though the light had been absorbed by the silver halide itself. Subsequently, the process of latent image formation is the same, regardless of whether light is absorbed directly or through the intervention of a sensitizing dye. The sensitometric properties of a dyed emulsion are not in general the same as those of the corresponding undyed emulsion. The differences are caused by secondary effects of the dye, such as may be concerned in development and in the distribution of sensitivity among the dyed grains, and not by any essential difference in the formation of the latent image.

Evidence for the identity of the photochemical process in the dye-sensitized emulsion with that in unsensitized material is given by measurements of photoconductivity. The photoconductivity of a dye-sensitized emulsion can be measured for different parts of the spectrum. In Figure 3 are shown the variation of photoconductivity, of photographic sensitivity, and of absorption with wavelength for three sensitizing dyes. The parallelism of the three effects is obvious.

The adsorption of dye to the silver halide can be followed by measuring the relation between the concentration of the dye solution in equilibrium with the dyed grains and the amount of dye adsorbed. Sensitizing dyes actually dye the silver halide, the dye molecules being firmly bound to the surface of the halide by intermolecular forces. In a common type of adsorption curve, the amount of adsorbed dye increases with the concentration of the dye added to the emulsion to a limiting value which represents complete coverage of the grain surface by a dye layer one molecule thick.

The limiting concentration of a dye which will saturate the silver halide and the surface area of the grains being known, the average area occupied per dye molecule can be computed. The typical sensitizing dyes are planar molecules, and in Figure 4 is shown one with an area of 140 square Ångstrom units, an edge-on area of 68 square Ångstrom units, and an end-on area of 35 square Ångstroms, as estimated

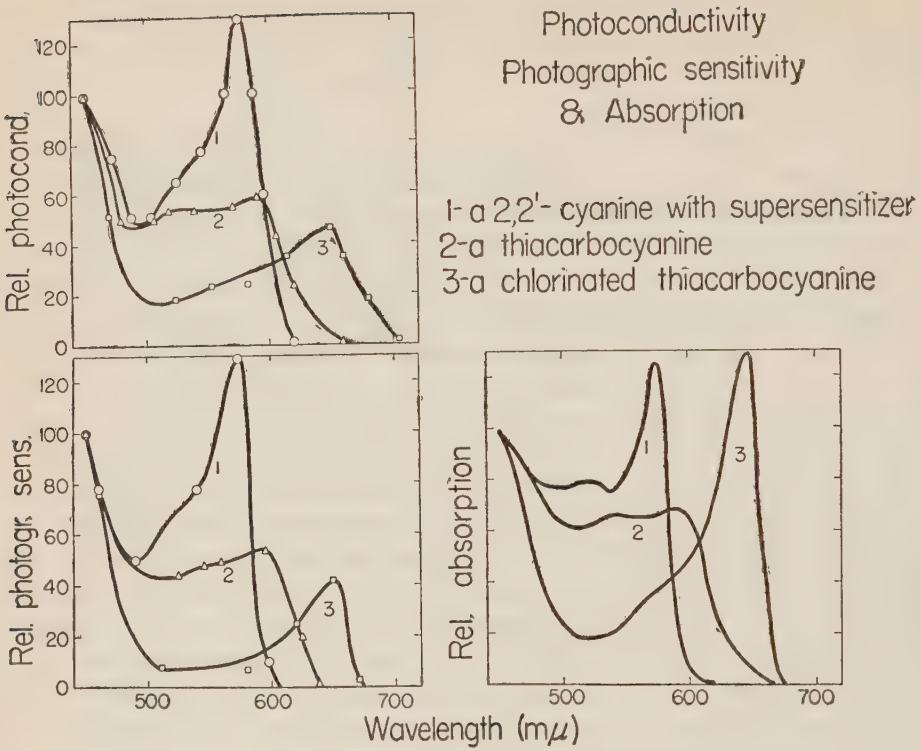


Figure 3

Photoconductivity, photographic sensitivity, and absorption of three sensitizing dyes.

from the assumed shape of the molecule. When this dye was adsorbed, it was found that each molecule occupied an area of 72 square Ångstroms, which is close to the calculated edge-on area. From this result, it is concluded that the dye molecules are stacked on the surface like rather thick books standing on a table with their long edges down.

The spectral absorption of sensitizing dyes varies in a complicated manner with the concentration. At low concentrations, the dye shows a strong characteristic maximum *M* (molecular band) with a weak secondary maximum. As the concentration increases, a new band develops at shorter wavelengths than the *M* band. This new band (*H*) may be ascribed to the formation of polymers and initially of dimers. Then, on further concentration, the dye develops an intense, sharp absorption maximum at a longer wavelength than the characteristic *M* band.

This new band (the *J*-band) is of a type reported independently by E. E. Jelley and by G. Scheibe, from whose work it appears that the molecules are stacked in a linear aggregate with the planes of the nuclei parallel and probably tilted to the axis. *J*-aggregation generally occurs much more readily on the grain than in solution.

The *J*-aggregates are not more efficient in sensitizing than the isolated molecules, but, since their absorption is unusually selective, the dyes which show them are often very valuable practical sensitizers, especially for the red and infrared.

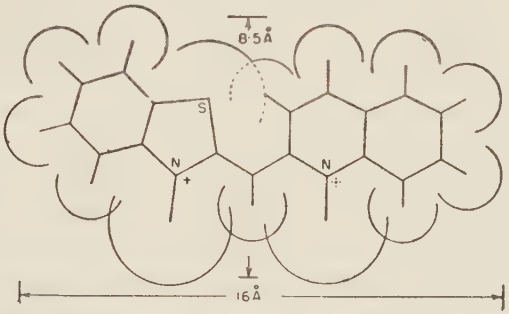


Figure 4

Geometry of a typical planar dye molecule.

Sensitization involves the transfer of energy absorbed by the sensitizer to the silver halide lattice. Optical sensitization by transfer is well known in simpler systems, and there is plenty of photographic evidence for this mechanism in silver halide emulsions. In the first place, optical sensitization acts almost independently of chemical sensitization. If an emulsion is made with an inert gelatin and one portion is left as a control and another treated with a chemical sensitizer, the second has several times the sensitivity of the first. If each of the portions is optically sensitized in the same way, both have practically the same distribution of optical sensitivity even though the absolute values of sensitivity at a given wavelength are quite different. Optical sensitization of an emulsion is usually the same for external and internal latent image formation; the distribution of latent image between external and internal image is the same for light absorbed by the dye as for light absorbed by the silver halide. The reciprocity failure of an emulsion may be changed by adding a sensitizing dye, but in the sensitized emulsion it is found to be independent of wavelength. As in other sensitized photochemical changes, the chemically effective transfer compares with a number of alternative processes, which can reduce the efficiency of transfer. The excited dye molecule, having absorbed light, may re-emit it as fluorescence; it may degrade the excitation energy to heat, returning to the normal electronic state by an internal radiationless transition; or it may transfer its energy to some component other than the silver halide; in any of these cases, the energy is lost photographically. All of these competing processes are observed with sensitizing dyes.

The action of a sensitizing dye is usually influenced by another adsorbed dye or compound. Generally, the effect is to diminish sensitization, but, as was mentioned earlier, certain dye combinations produce a considerable increase in the sensitization, which is called *supersensitization.* A consistent explanation of supersensitization can

be given in terms of the hypothesis of energy flow from molecule to molecule in the adsorbed dye layer. It will be recalled that the results on adsorption suggested that the dye molecules are held edge-on to the grain, with the planes of the nuclei of adjacent molecules projecting from the surface parallel to each other.

When energy is absorbed in such a system, it may travel through the pack of molecules until it is dissipated as heat or until it enters the silver bromide and forms a latent image. The probability of the energy being transferred to the silver bromide is greatly enhanced if there is in the layer a discontinuity which will slow down the rate of propagation without stopping it altogether. A dye which acts as a supersensitizer must therefore be sufficiently similar to the sensitizer to be incorporated in a kind of solid solution in the two-dimensional crystal constituting the dye layer but sufficiently dissimilar to be a perturbation. Hence it is likely to be a dye of a somewhat different class from the sensitizer. It must also not acquire the energy itself and lose it by fluorescence, internal degradation, or photochemical decomposition.

Dyes have been found which act not at supersensitizers but as antisensitizers, and it is believed that they act by capturing the energy at the discontinuity, as the supersensitizing dyes do, but the energy is degraded internally instead of being transferred to the silver bromide.

TEACHING EXPERIMENTS ON THE CORRELATIONS BETWEEN THE DENSITY CURVES OF PHOTOGRAPHIC LAYERS OF NEGATIVES AND POSITIVES

J. EGGERT

Photographic Institute of the Federal Institute of Technology, Zurich

It is most peculiar that photography was practised already for half a century for practical and scientific purposes before Hurter and Driffield created its scientific foundation — sensitometry. Only after three more decades had elapsed, did E. Goldberg use this basis for his comprehensive and famous investigation of the structure of the photographic image[1]. Though these things are today the assured property of science, it will be observed that instruction in this subject is as yet not solidly founded, particular regarding Goldberg's findings. Therefore, we here in Zurich undertook, years ago, to work out a practical case, aiming at a clear demonstration of these things and to show, above all, the limitations of photography.

For this purpose, we used a model consisting of a pictorial and a systematic part. Figure 1 shows the arrangement, but the following has to be considered: as the transparency (mountain scenery) is strongly illuminated, whilst the book, vase and picture receive light from weak lamps only, the object-range of the model, which is the relation between the luminance of the most strongly "shining" part of the object and that of the weakest radiating part, is well above 1000. With the aid of the adjoining transparency, in which a range of films of well defined densities is illuminated from behind by a light of even luminosity, the luminance of individual parts of the object can easily be estimated. The table at the left, showing numerical values of the relative luminance J of the steps, the transparency T and the density D of the films producing them, gives the quantitative basis of the experiment. In preparing the pattern for the reproduction of Figure 1, conditions of illumination had changed so that the object range of model and scale was considerably reduced. This was done in order to give an approximate picture of the object.

The model was then photographed in an otherwise darkened room with a camera fixed on a tripod, on negative plates of the size 9×12 cm, with constant diaphragm and graduated exposure time (0.25; 1; 4; 16; 256; 1024 sec). The 7 negatives thus obtained are shown in Figure 2 (top), mounted on glass and illuminated from

Received May 25, 1956.

behind. Moreover, Figure 2 shows the density curve of the emulsion (Gevaert Superchrom) according to the respective conditions of development. It was purposely kept very flat ($\gamma = 0.53$). Below it are those parts of the density curve which were used in the production of the 7 negatives (the additional stepwedge of densities is given for reference only and does not conform to the values of the illuminated stepwedge of Figure 1, which, too, rises much slower than in reality). The relative luminances J of the individually numbered fields of the scale, as per table of Figure 1, would result in corresponding densities according to the density curves of Figure 2, provided the objective and camera walls issue no scattered light. The value of the latter can easily be determined with Goldberg's device[2] in the case under consideration; a pronounced correction of the J-values results, which in our case proved to be so unimportant that it could be ignored.

Figure 3 illustrates the printing of the negatives on contrast paper (Tellko-Fribourg). The 7 negatives (vertical lines) were printed with 7 different exposure times under identical conditions of light and development (horizontal lines), and the resulting 49 pictures were examined as to quality. As expected, none of them is a satisfactory reproduction of the model. Either the picture in the interrupted frame suffices for the upper part of the model, or the one in the dotted frame suffices for the lower part of the model; although the fully framed picture shows the entire object, all its parts, particularly the lights, appear in such a weak contrast that this compromise must be considered very unsatisfactory. The reason for this failure is, of course, the fact that the enormous object range of the model puts demands which the photographic process cannot meet. As Goldberg was probably the first to recognize, the photographic image on paper has the very limited density range from 1.6 to a maximum of 1.8 (numerical value from 1:40 to 1:60) and not that of the model, which is nearby double, namely 3.3 (numerical value 1:2000).

The right hand part of Figure 3 gives the facts in the form of graphs. It is intended to show how negative and positive materials build the picture. Goldberg postulated[3] that for a faithful reproduction of contrasts the gradients n_N of the negative material and n_P of the positive materials be adjusted so that the product

$$n_N \cdot n_P = 1, \tag{1}$$

results, which is called the Goldberg condition. We modify this classic equation in order to generalize it for any printing process. First, we consider the "copy curve", in which the logarithms of the luminance of the original are shown opposite the densities of the print and therefore definitely show the functional connection of original and reproduction[4]. This curve runs opposed to the density curves of the negative and positive materials and possesses negative inclination values n_{Cop}. In general, equation (1) with respect to corresponding points of the three curves will be of the form:

$$n_N \cdot n_P = -n_{\mathrm{Cop}}. \tag{2}$$

No	J	T	D
12	2048	1	0
11	1024	0.5	0.3
10	512	0.25	0.6
9	256	0.125	0.9
8	128	0.063	1.2
7	64	0.031	1.5
6	32	0.016	1.8
5	16	0.008	2.1
4	8	0.004	2.4
3	4	0.002	2.7
2	2	0.001	3.0
1	1	0.0005	3.3

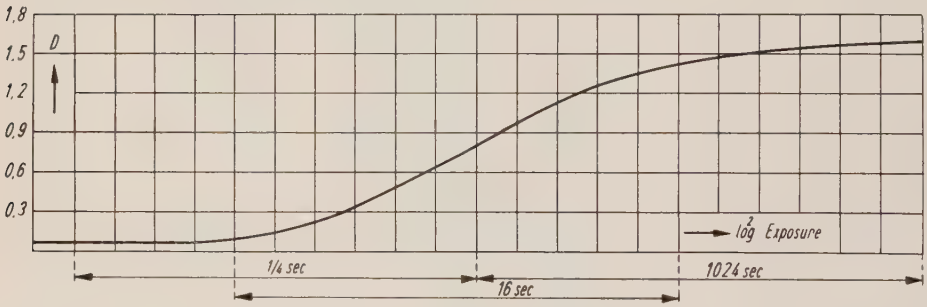
Figure 1

The model used for the experiment (strongly illuminated transparency of a mountain scenery, a few weakly illuminated objects, and transparency of a stepwedge); object range of the original 1:2048; the object range was much reduced for the above reproduction.

For the straight parts of the three density curves applies the formula

$$\gamma_N \cdot \gamma_P = -\gamma_{\mathrm{Cop}}. \tag{3}$$

To construct the copy curve from the two others, one proceeds in the following way, as can be seen from the diagrams of Figure 3: the individual density values



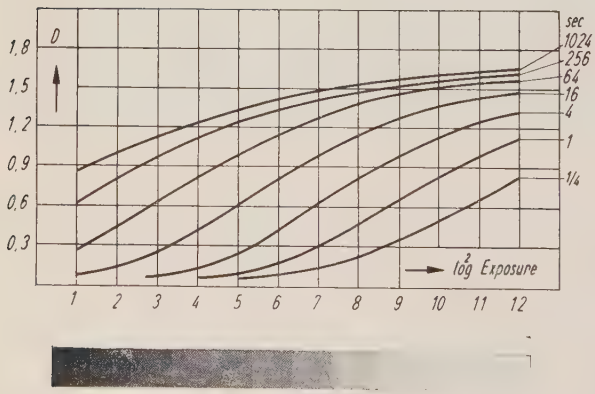
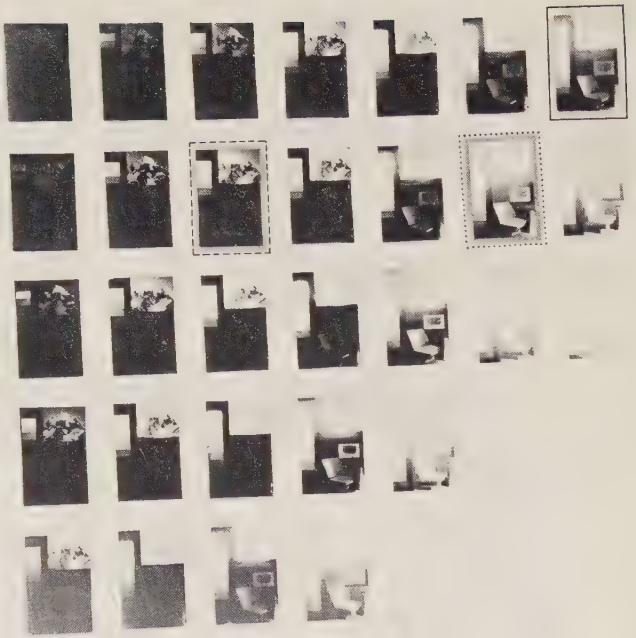


Figure 2

TOP: The 7 negatives of the photographs of the model.

CENTRE: Density curve of the negative material.

BOTTOM: Parts of the density curve used for the 7 negatives.



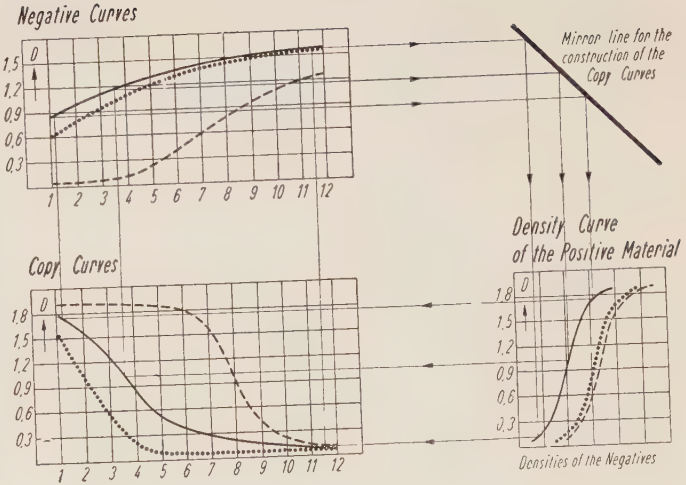


Figure 3

LEFT: 49 prints of the 7 negatives on contrast paper; vertical columns of same negatives, horizontal rows of some printing times increasing upwards with the factor 2.

TOP: Construction of the copy curve. Frames of prints and style of curves correspond. The construction lines belong to the fully drawn curves which correspond to the fully framed image on the left.

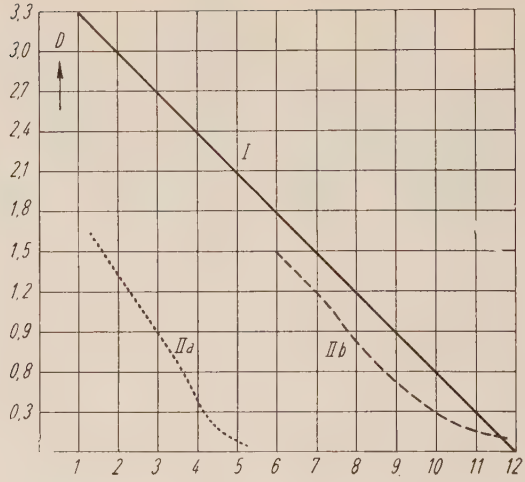


Figure 4

Mounting of the reproduction of the model (BOTTOM LEFT) from two parts of pictures (TOP LEFT); adjoining are the copy curves (for particulars see text).

of the negative are transferred to the abscissa of the density curve of the printing paper by mirroring in a straight line running at an angle of 45° to the coordinates. The abscissa has to be placed in such a manner that on printing (or enlarging) the greatest density of the negative results in the smallest density of the paper, and vice versa. The wanted copy curve is then obtained at the intersections of vertical lines from the values of luminance of the original with horizontal lines drawn from the corresponding points on the density curve of the positive material. This construction has been carried out in respect of the three previously selected prints of Figure 3, and the construction lines pertaining to the 3 curves drawn are added. It can be seen that the three copy curves correspond to the similarly marked images, especially with regard to the steps of the wedge. The interrupted curve shows graduations rich in contrasts only in the lights (fields 7 to 12), the dotted curve only in the shadows (fields 1 to 5); the solid curve indeed covers the entire field, but with flat graduations, particularly in the lights (fields 5—12), and corresponds exactly to the fully framed print of the model.

It is an obvious and sometimes done combination to "mount" the images framed with dots and dashes, as shown in Figure 4. One look at the density step lays bare this trick, because the straight (ideal) copy curve I in Figure 4 is replaced by the two connecting part curves IIa (shadow) and IIb (light). If need be, the model can be reproduced from the negative of the fully framed image, if the print is not made on paper but on contrast transparency material, whose density range exceeds by far that of paper.

Experiments in printing similar to that of Figure 3 have been carried out on printing paper of other grades. Their rendition is superfluous as they do not differ in principle.

However, the practical limits for the reproduction of detail in light and shadow, in the meaning of Goldberg[5], can easily be taken from the curves of Figure 3 and compared with the corresponding parts of the image. This is an exercise valuable for the understanding of the procedure, but we can renounce it here, the more so because Goldberg's classic explanations open the way to it.

REFERENCES

1. Goldberg, E., *Der Aufbau des photographischen Bildes*, W. Knapp, Halle (Saale), 1st ed. 1922, 2nd ed. 1925. A thorough appreciation of this valuable work, whose rich and classic contents cannot be discussed here, is found in: Weigert, F., 1922, Ein photographisches und optisches Standardwerk, *Naturwiss.*, **10**, 861.
2. *Ibid.*, 2nd ed., p. 22.
3. *Ibid.*, 2nd ed., p. 64.
4. The first but less clear exposition of this "reproduction curve" can be traced back to Jones, L. A., 1920, "On the theory of tone reproduction, with a graphic method for the solution of problems", *J. Franklin Inst.*, **190**, 39. Indeed, this independent and simultaneous research deals with problems which are also discussed by Goldberg; thus Jones, too, postulates equation (1).
5. Goldberg, E., *loc. cit.*, p, 74.

LES INSTRUMENTS D'OPTIQUE ET LA LUMIERE PARASITE

A. Arnulf
Institut d'Optique, Paris

ABSTRACT

This paper is related to the application of Goldberg's studies concerning stray light to the optical instruments which are built in the Institut d'Optique, Paris.

The causes of the stray light and its effect on the perceptions are given. Methods and apparatus used for checking the visual or the photographic instruments are described.

En 1936 paraissait en France la traduction du livre du Professeur E. Goldberg, *La Formation de l'Image Photographique*[1]. Les résultats de cette étude, qu'il s'agisse des systèmes optiques, des propriétés des émulsions, des méthodes d'investigation et de contrôle, sont maintenant classiques; mais, sur plusieurs points, leur portée dépasse le domaine de la photographie. En particulier, nous retiendrons deux conclusions qui s'appliquent à tous les types d'instruments.

(1) La perception ou la reproduction des contrastes faibles, spécialement en présence de grandes oppositions de lumière, est un des éléments fondamentaux caractérisant la qualité d'un instrument d'observation.

(2) Il résulte de la proposition précédente le fait que la lumière parasite est un des défauts les plus nuisibles de ces mêmes instruments, et cela d'autant plus que les contrastes sont plus faibles et les oppositions de lumière plus violentes.

Goldberg a eu le mérite de mettre en évidence l'importance de la lumière parasite; et encore ses effets sont-ils beaucoup plus notables dans les instruments visuels que dans les instruments photographiques, ces derniers ne demandant pas, le plus souvent, une image aussi bonne.

L'étude de la lumière parasite et de son influence sur la perception a été, durant plusieurs années, un des centres d'intérêt de l'Institut d'Optique de Paris. Je pense que le meilleur hommage que nous puissions rendre au Professeur Goldberg est d'exposer ici comment nous avons mis ses idées en oeuvre.

CAUSES ET EVALUATION DE LA LUMIERE PARASITE

Goldberg signale 3 causes principales de lumière parasite:

(1) Les réflexions d'ordre pair sur les surfaces air-verre. Le flux incident sur une telle surface est réfléchi par elle dans la direction de l'objet, puis réfléchie de nouveau, cette fois dans la direction de l'image, par toutes les surfaces qu'il rencontre; on superpose ainsi à l'image un voile de lumière parasite.

L'uniformité et l'étendue de ce voile dépendent de la distribution des luminances dans l'objet, de la position des images réfléchies d'ordre pair et de la

Received August 17, 1956.

position des pupilles et des lucarnes; elles sont donc essentiellement variables avec les divers types d'instruments.

(2) La diffusion par les surfaces, qui peut provenir d'un polissage incomplet ou d'un nettoyage insuffisant. La lumière tombant sur une particule diffusante est détournée de sa route par réfraction et par diffraction et répartie dans l'espace suivant une indicatrice qui dépend de la forme et de la dimension de la particule.

(3) La diffusion et la réflexion par les parois internes de la monture des lentilles. Pour les objectifs photographiques, et plus généralement pour les systèmes qui forment une image réelle sur une couche sensible, cette lumière parasite s'ajoute simplement à celle qui provient des lentilles. Dans les instruments visuels par contre, elle se répartit autour de l'anneau oculaire et pourra ne pas être gênante, tant que ce dernier a un diamètre plus grand que celui de la pupille de l'oeil et que cette dernière reste centrée.

Depuis le livre de Goldberg, un certain nombre de faits nouveaux se sont produits.

Les traitements de surfaces antiréfléchissants permettent à la fois de diminuer les pertes de lumière et la lumière parasite. C'est là un progrès considérable, qui se traduit essentiellement par une amélioration des performances instrumentales, d'autant plus marquée que les contrastes sont plus faibles. Les traitements antiréfléchissants n'annulent pas la réflexion; ils ne peuvent donc pas supprimer complétement la lumière parasite. Mais, lorsqu'on en calcule le résidu, et qu'on le compare aux valeurs mesurées, on constate que ces dernières sont toujours notablement plus fortes. Ces écarts peuvent alors être attribués à 3 causes.

(a) La diffusion par les couches minces antiréfléchissantes; une diffusion, assez faible pour échapper complétement à l'observation banale peut, non seulement neutraliser le bon effet du traitement, mais encore aggraver la situation. Il arrive que la transparence de l'objectif soit augmentée, tandis que les contrastes de l'image sont rendus plus faibles. Ces cas sont fréquents, et les constructeurs les plus réputés n'en sont pas exempts; c'est une des plus grandes difficultés de la fabrication des couches minces que de les obtenir sans aucune diffusion, et cela exige des précautions rigoureuses et un contrôle constant.

(b) Une fraction importante des résidus de lumière parasite provient des micro-défauts de forme des surfaces, révélée par B. Lyot, par l'élégante méthode de contraste de phase qu'il avait mise au point pour le contrôle des lentilles du coronographe qui, on le sait, est l'instrument destiné à l'observation de la couronne solaire; il s'agissait là d'un cas extrême, puisqu'on demandait au flux parasite de ne pas dépasser quelques millionièmes du flux provenant du soleil. Cela nécessitait une qualité de poli très difficile à réaliser et qu'il est tout à fait impossible d'obtenir dans le travail industriel[2].

L'amélioration obtenue par le perfectionnement du poli des surfaces n'aurait vraisemblablement pas un grand intérêt pour les objectifs photographiques, parce que les émulsions courantes "mangent" les contrastes faibles. Par contre,

elle permet des gains sensible des performances des instruments visuels dans les cas, peut être rares, mais souvent très importants en pratique, où il faut voir à tout prix des détails dans de violentes oppositions de lumière.

(c) Une dernière fraction de cette lumière parasite provient de la nature même de la tache de diffraction, qu'on assimile souvent au disque central d'Airy, alors que 17% environ du flux incident se répartit dans les anneaux extérieurs et se comportent, en fait, comme de la lumière parasite. C'est ce qu'ont calculé et mesuré Francon et Cagnet[3], en utilisant le dispositif expérimental de Goldberg[4]. Cette lumière parasite peut atteindre des valeurs importantes lorsque les dimensions angulaires de l'objet sont petites.

La suppression des anneaux entourant le disque d'Airy aurait donc pour effet d'augmenter les contrastes. Cela est possible grâce aux techniques dites d'"apodisation", dues à Couder et à Jaquinot, et sur lesquelles on trouvera renseignements et documentation dans un travail de Brigitte Dossier[5].

TABLEAU I

Valeurs du rapport δ de la luminance parasite à la luminance de l'image, mesuré par la méthode de Goldberg

(*A*) *Jumelles à prismes*

Nature	7 × 50	20 × 70	4 × 20	6 × 30 droite	6 × 30 gauche
Nombre de surfaces traitées	4		10		
Nombre de surfaces non traitées	2	12	2	12	12
δ	1.5%	3.3%	3.2%	41%	25%
Remarques			Traitement diffusant	jumelle voilée	jumelle voilée

(*B*) *Objectifs photographiques*

Nature	f = 85mm F/2		f = 50mm F/2		f = 50 mm F/3.5	f = 145mm F/4.5
Nombre de surfaces traitées	6		5			
Nombre de surfaces non traitées	6		6	1	6	4
δ	10.7%	2.9%	7.2%	3.0%	2.6%	1.3%
Remarques	Cet objectif a été mesuré avant et après traitements antiréfléchissants		Mesures avant et après traitement. Face avant, attaquée, non traitée			

A. Saez[6] a montré, par des mesures directes, que l'apodisation améliore sensiblement les contrastes, et par conséquent le pouvoir séparateur sur les objets de faibles contrastes.

La combinaison de l'apodisation, de l'amélioration de la qualité du poli des surfaces et la suppression de toute diffusion dans les traitements antiréfléchissants, permettrait de réaliser des instruments pratiquement exempt de lumière parasite, ainsi que l'a fait Lyot dans son coronographe. Mais un tel instrument pourrait-il conserver ses qualités pendant un temps suffisant? La lumière parasite dépend énormément de l'état des surfaces qui, à la longue finit par s'altérer sous l'action de l'atmosphère.

Le tableau I donne quelques valeurs typiques de la lumière parasite pour divers types d'instruments. On remarquera, d'une part, l'effet de la diffusion par les surfaces, d'autre part, l'amélioration importante obtenue par les traitements antiréflèchissants, lorsque les couches sont exemptes de diffusion.

EFFET DE LA LUMIERE PARASITE SUR LA PERCEPTION

Ayant eu à m'occuper de la construction de périscopes de sous-marins qui, avec leur grand nombre de lentilles, sont les instruments d'élection de la lumière parasite, je me suis heurté dès le début au fait paradoxal suivant: parmi des instruments de construction différentes, disons A et B, les uns, A, se révélaient nettement supérieurs à B lorsqu'on les contrôlait au laboratoire (mesure de la limite de résolution en utilisant un test de Foucault de fort contraste placé sur un fond noir), tandis que, utilisés dans la nature, les instruments B possédaient une supériorité si écrasante que les utilisateurs n'en voulaient pas d'autres.

Il fallut bientôt se rendre à cette évidence que la méthode de contrôle de laboratoire ne faisait intervenir que les défauts entraînant un élargissement de l'image, par exemple les défauts de mise au point ou l'astigmatisme, et éliminait à peu près totalement l'effet nuisible de la lumière parasite, qu'elle provienne du voile, ou même des aberrations sphérique ou chromatique.

Des résultats utilisables n'ont pu être obtenus qu'aux conditions suivantes:

(a) Faire intervenir la totalité de la lumière parasite susceptible de se former dans l'instrument, et, pour cela s'arranger pour que le flux incident couvre la totalité de l'hémisphère antérieur de l'instrument.

(b) Mesurer les limites de résolution pour plusieurs contrastes, échelonnés entre le maximum et la plus petite valeur perceptible.

Les comparaisons effectuées entre les instruments A et B précédents en utilisant cette méthode ont montré une supériorité évidente de B pour les faibles contrastes; c'est donc eux qui déterminent, pour les usagers, la qualité de l'instrument.

En fait les mesures de limites de résolution, même effectuées par des observateurs ayant une vue dite normale, peuvent varier d'un individu à l'autre. On obtient des résultats bien plus constants en formant le rapport E de la limite de résolution s de l'oeil nu, pris avec le même diamètre pupillaire que lorsqu'il est associé à

l'instrument, et la limite correspondante, s', de l'oeil observant le champ image de l'instrument, l'objet étant évidemment identique dans les deux cas.

$$E = s/s' = \text{efficacité.}$$

L'efficacité représente la perte de perception causée à l'oeil par la présence de l'instrument. Si $E = 1$, l'instrument ne change pas cette perception, il est visuellement parfait; si $E < 1$, l'oeil voit moins de détails dans le champ image que directement; l'instrument abîme la vision, il est imparfait. L'efficacité revient en fait à faire les rapports des limites de résolution d'un instrument parfait et d'un instrument imparfait; elle exprime donc la qualité de l'instrument.

Un instrument sera complétement défini de ce point de vue par la courbe donnant les valeurs de l'efficacité en fonction du contraste, ou plus simplement, les valeurs de l'éfficacité pour 2 ou 3 contrastes convenablement choisis. Lorsqu'on extrapole la courbe précédente jusqu'à l'efficacité zéro, on obtient le seuil de contraste perceptible par l'oeil associé à l'instrument.

Le cas des instruments photographiques se traite d'une manière analogue, mais fait intervenir les limites de résolution linéaires de l'oeil associé à l'émulsion.

Nous pouvons évaluer les pertes de contraste dues à la lumière parasite. Si on sait, d'autre part, comment varie la perception en fonction du contraste on peut répondre, d'une manière au moins approximative à la question posée sur l'effet produit par la lumière parasite sur la perception.

Cherchons à évaluer les pertes de contraste que l'on pourra rencontrer. Goldberg a fait de ce point de vue une étude étendue des images photographiques. Nous nous bornerons à deux cas extrêmes, en simplifiant au maximum. Le premier est celui d'un objet dont les luminances sont peu variables, par exemple, un paysage lointain estompé par la brume, ou au laboratoire, un test objet se détachant sur un fond uniforme de luminance analogue.

Adoptons la définition du contraste utilisée pour les études visuelles, soit:

$$C_0 = F - O/F,$$

F étant la luminance du fond, O, celle du test.

Supposons que la luminance du voile de lumière parasite soit uniforme, et représente une fraction δ de la luminance de l'image du fond. On voit facilement que, C étant le nouveau contraste en présence de lumière parasite, on a:

$$C = C_0/1 + \delta \quad ; \quad C_0 - C/C = \delta$$

Si nous prenons les valeurs de δ mesurées dans les bons instruments, telles qu'elles figurent dans le tableau I, la variation relative du contraste est très faible; même pour de fortes valeurs de δ (par exemple $\delta = 0.1$) la perception est peu changée. Il ne faut pas cependant perdre de vue que, lorsqu'un détail ou une plage étendue sont à la limite du contraste perceptible, sans voile, ils disparaitront en présence d'un voile, même peu intense.

Le deuxième cas correspond à de violentes oppositions de lumière; par exemple, lorsqu'on cherche à voir des détails sur un avion sombre se détachant sur un ciel clair, ou à la lisière d'un bois situé au voisinage de l'horizon, par temps sans brume, etc.

Au laboratoire, ce cas peut être réalisé par un objet simple, composé par exemple de deux plages juxtaposées de contraste C_0, se détachant sur un fond étendu de luminance plus forte. A étant la luminance de la plage la plus claire de l'objet, F, celle du fond, posons:

$$F/A = K, \text{ avec } K > 1$$

On trouve, C étant le contraste résultant lorsqu'on fait intervenir la lumière parasite,

$$C = C_0/1 + K\,\delta$$

L'abaissement de la luminance de l'objet par rapport à celle du fond a pour effet de multiplier δ par K.

$$\begin{aligned} \text{Pour } \quad K &= 50, \ \delta = 0.1, \ \text{ on a: } \ C = C_0/6 \\ K &= 1 \ , \ \delta = 0.1, \ \text{ on a: } \ C = C_0/1.1. \end{aligned}$$

La figure 1 donne, en fonction du contraste image C_0 sans lumière parasite, les rapports E des limites de résolution s, sans lumière parasite, à s', avec lumière parasite, pour diverses valeurs de:

$$Q = 1/1 + K\delta.$$

On voit combien est important l'effet de la lumière parasite pour les contrastes faibles. Il convient du reste de noter qu'un observateur normal apprécie très nettement une baisse d'acuité correspondant à $E = 0.9$.

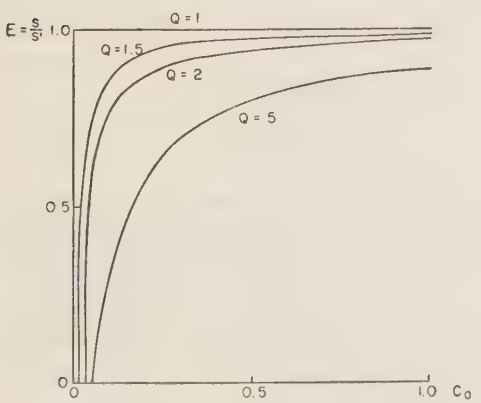


Figure 1

Effet de la lumière parasite sur l'efficacité (test de résolution).

On obtiendrait des courbes analogues avec des tests de perception tels que lignes ou disques, mais on verrait alors que l'effet nuisible de la lumière parasite pour les contrastes forts est plus important que pour les tests de résolution.

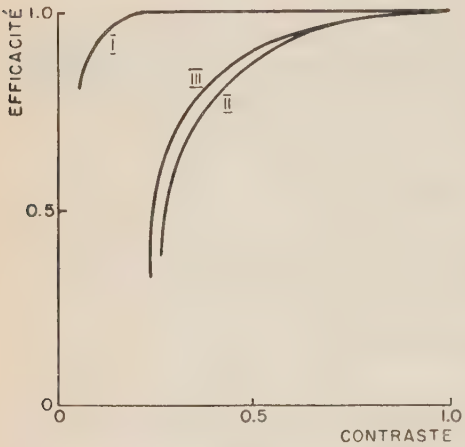


Figure 2

Courbes expérimentales donnant l'efficacité en fonction du contraste.

I. Lunette d'excellente qualité. Effet d'une petite quantité de lumière parasite.

II. La même qu'en I, mais l'objectif est remplacé par une lentille plan convexe au minimum d'aberration sphèrique (effet dominant du chromatisme).

III. Comme en I, mais l'objectif est retourné (effet de l'aberration sphèrique).

Les courbes expérimentales obtenues avec les instruments visuels affectent les formes indiquées (fig. 2); par contre, les résultats sont fort différents pour les objectifs photographiques travaillant avec les émulsions courantes. Les courbes ont alors l'allure de la figure 3.[7] Cela s'explique par le fait que ces émulsions reproduisent si mal les contrastes faibles que la perte de contraste due aux défauts instrumentaux

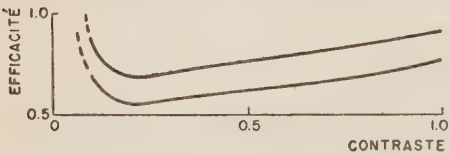


Figure 3

Deux objectifs, type Tessar, ouverts à f/6.3. Courbes expérimentales de l'efficacité en fonction du contraste.

n'a pratiquement aucun effet sur leur reproduction; les objectifs se comportent comme s'ils étaient parfaits lorsque les contrastes sont faibles.

APPAREILS EN SERVICE A L'INSTITUT D'OPTIQUE

La nécessité de faire intervenir la lumière parasite pour la mesure des performances instrumentales nous a conduit à jumeler cette installation avec celle destiné à la photométrie. Dans ce dernier cas, le principe de la mesure est celui mis en oeuvre par Goldberg, dans lequel on admet que la luminance parasite est celle de l'image d'un trou parfaitement noir, placé au milieu d'un champ uniformément éclairé, et dont les dimensions sont assez petites pour ne pas perturber sensiblement le flux incident; lorsqu'il s'agit de mesurer des limites de résolution ou des efficacités, on fait apparaître le test à la place de la plage noire, sans rien changer à l'éclairement du champ.

Les études et les contrôles d'instruments visuels sont pour nous de beaucoup les plus fréquents; il importe donc que la plage noire ou le test qu'on lui substitue soient situés assez loin de l'instrument pour pouvoir être mis au point; la présence de toute pièce optique auxiliaire doit être absolument prohibée entre le test et l'instrument, car on introduit ainsi des quantités non négligeables de lumière parasite, ce qui entraine des corrections peu sûres et complique la méthode.

En fait, nous avons adopté la distance maximum dont nous disposions, soit 40 mètres.

CONDITIONS EXPERIMENTALES STANDARD

Nous avons vu que la proportion de lumière parasite varie avec la répartition des luminances et la répartition des sources dans le champ et même en dehors du champ. Sa mesure ne peut avoir de sens que si l'on définit les conditions expérimentales. La même remarque s'applique au cas des mesures d'efficacité.

Nous avons adopté les dispositions standard suivantes: deux cas ont été prévus, suffisants pour les besoins de la pratique.

(a) Il n'y a aucune source éblouissante dans le champ de l'instrument: par convention l'objet est une surface indéfinie, de luminance uniforme comprise entre 0.1 et 0.01 b/cm^2, ces valeurs étant imposées pour assurer une bonne photométrie visuelle et de bonnes mesures de limites de résolution.

(b) Le deuxième cas est celui où existe une source éblouissante dans le champ ou son voisinage immédiat (par exemple, le soleil, la lune, un projecteur).

Nous convenons alors d'utiliser une source se présentant sous la forme d'un disque qui sous tend, depuis la pupille d'entrée de l'instrument, un angle de 0.01 radian; cette source peut occuper les positions suivantes: la moitié du rayon du champ, tangente *extérieurement* au champ, de manière qu'aucune partie ne soit vue directement; enfin, le centre du disque est à une distance angulaire du bord du champ égale au demi rayon de ce champ.

L'effet produit en un point donné du champ par la combinaison d'une surface de luminance uniforme et d'une source éblouissante est la somme des lumières parasites correspondantes; il est souvent commode d'étudier les deux cas séparément, en opérant sur un fond entièrement noir pour la source éblouissante; il est facile, connaissant son intensité, d'en déduire la lumière parasite produite par une source quelconque d'intensité connue.

REALISATION D'UNE PLAGE UNIFORME D'ETENDUE INDEFINIE

On utilise le dispositif de la figure 4. L'instrument à étudier J, est porté, ainsi que le photomètre, sur un support approprié et vise horizontalement; il est engagé à l'intérieur d'une sphère diffusante S, peinte en blanc intérieurement, et éclairée au moyen de deux lampes L, munies chacune d'un écran qui les masque à l'objectif de l'instrument J. Ce dernier vise à travers une ouverture circulaire T, de diamètre réglable, et située sur le diamètre horizontal de la sphère qui passe par l'objectif.

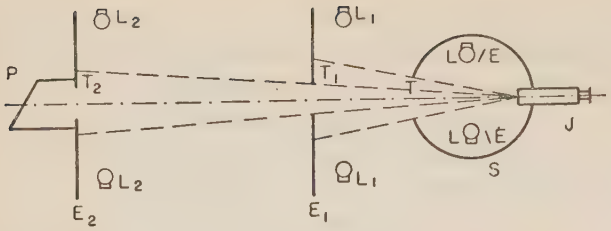


Figure 4

Deux écrans circulaires, E1 et E2, éclairés uniformément par des couronnes de lampes L1 et L2, permettent d'éclairer la totalité du champ. Ils portent des ouvertures circulaires T1 et T2 de diamètres réglables, T2, munie d'un piège à lumière P, constituant la plage noire objet des mesures de lumière parasite, et T1 ayant le diamètre voulu pour laisser passer intégralement les faisceaux s'appuyant sur les bords de T2. La luminance de l'intérieur de la sphère et de la surface des écrans est réglée à l'égalité, en faisant varier la distance des couronnes de lampes L1 et L2 aux écrans E1 et E2.

REALISATION DE LA SOURCE BRILLANTE

On projette (fig. 5.) sur la pupille d'entrée O de l'instrument étudié l'image S′ d'une source uniforme S de luminance connue, au moyen d'une lentille L. Le diamètre de L est choisi de manière que l'angle qu'il sous tend depuis la pupille d'entrée O soit

Figure 5

égal à 0.01 radian; la longueur focale de L est choisie de manière que l'image S′ couvre complétement la pupille O. La surface de la lentille L joue alors le rôle d'un objet de luminance $B = B_0T$, B_0 étant la luminance de la source et T le facteur de transmission de la lentille. On prend pour S une lampe à ruban de tungstène, ou une lampe pointolite, ou éventuellement un arc électrique.

PHOTOMETRE, SUPPORTS DE L'INSTRUMENT ET DU PHOTOMETRE

Voici le dispositif utilisé pour les instruments visuels. Le schéma du photomètre est représenté sur la figure 6. Il est du type "à comparaison" et on trouvera la description d'un modèle du même genre dans l'ouvrage cité de Goldberg. On compare, au moyen d'un cube de Lummer, les diverses luminances du champ image de l'instrument à une luminance caractéristique de l'objet, par exemple celle du fond, B_0. On détermine en général les rapports:

$$T_a = B_a/B_0 \; ; \; P = B_p/B_0 \, ,$$

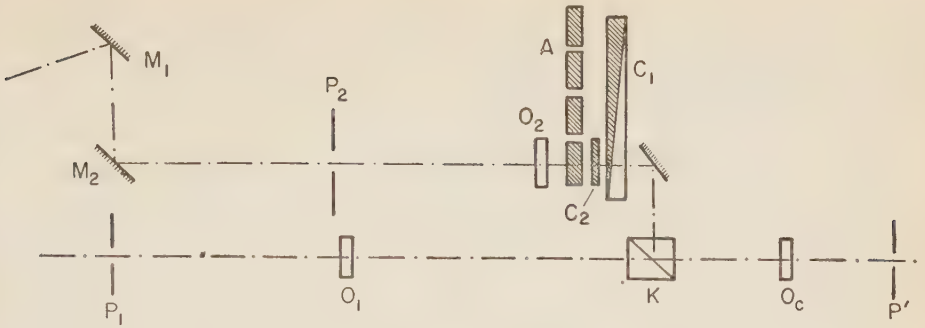


Figure 6

B_a étant la luminance correspondant à B_0 vu à travers l'instrument, B_p étant la luminance du trou noir (luminance parasite).

On en déduit le facteur de transmission effectif T, c'est-à-dire corrigé de la lumière parasite, soit:

$$T = B_a - B_p/B_0 = T_a - P$$

et le pourcentage de la lumière parasite voilant l'image,

$$\delta = B_p/B_a - B_p = P/T.$$

Le trajet de mesure vise à travers l'instrument, et comprend: le diaphragme variable P1, qui définit la pupille d'entrée, et l'objectif O1 qui projette l'image du champ sur le cube de Lummer K. Le trajet de référence comprend d'abord le miroir M1, mobile autour d'un axe horizontal et d'un axe vertical, puis le miroir M2 fixe, renvoyant le faisceau incident suivant l'horizontale. Par rotation de M1 on peut viser un point situé dans des zones étendues au-dessus et au-dessous du plan horizontal. La pupille d'entrée est un iris identique à P1, et se situe en P2; elle est suivie de l'objectif O2, qui projette l'image de la plage de référence sur le cube K, par l'intermédiaire du miroir M3. Le dispositif de gradation comprend un coin photométrique C1 en verre neutre, avec son contre-coin C2, et une serie de surcharges interchangeables A, permettant de mesurer des luminances variables dans un rapport 10^4. Oc est l'oculaire, et P′ la pupille de sortie.

L'ensemble de l'instrument et du photomètre est porté par un petit banc optique, B, pouvant tourner autour de l'axe vertical A1, placé à l'intérieur de la sphère S (fig. 7). L'instrument J repose sur un support réglable P1, permettant d'amener sa pupille d'entrée sur l'axe de rotation A1 et de régler son axe optique à peu près parallèle à l'axe du banc B. Le photomètre Ph est porté par un support reposant sur le banc B, et qui pivote autour de l'axe vertical A2, d'angles qu'on lit sur une division D et qui représentent les angles de champ image. Par construction, la pupille d'entrée du photomètre coincide avec l'axe A2, et l'ensemble du support et de l'axe A2 peut être déplacé de manière à amener la pupille d'entrée du photomètre sur

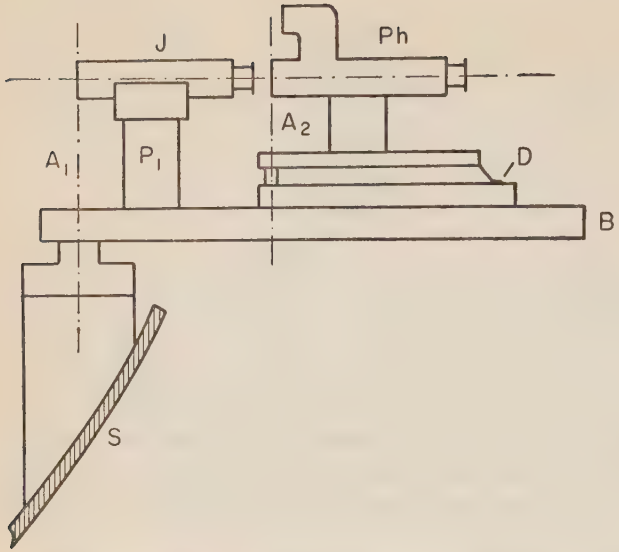


Figure 7

l'anneau oculaire de l'instrument J. La rotation du banc B déplace l'objet dans le champ objet; on pointe l'image par coïncidence avec la plage photométrique de référence, et on mesure sa position angulaire dans le champ image sur le cercle divisé D.

Dans le cas de la photométrie d'un objectif, on opère un peu différement (fig. 8). Le photomètre est muni d'une lentille auxiliaire L, qui projette l'image de la pupille

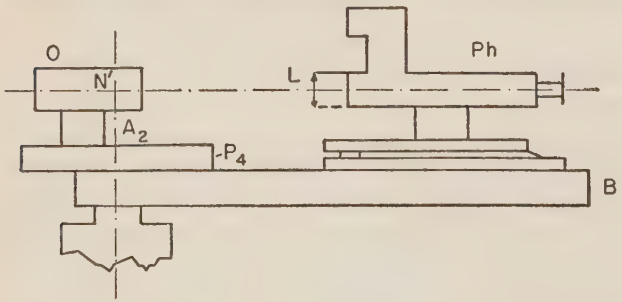


Figure 8

de sortie de l'objectif dans la pupille d'entrée du photomètre; elle joue en fait le rôle de l'oculaire d'un instrument d'observation visuelle. L'objectif O est placé de manière que son point nodal image N′ soit sur l'axe A1. Il peut tourner autour de cet axe, les rotations étant repérées sur le cercle divisé P4. Le banc B est immobilisé, son axe étant parallèle à la ligne d'observation; le photomètre est placé sur cet axe, il est mobile longitudinalement le long du banc B, de manière à assurer la mise au point correcte en toutes les régions du champ. Lorsqu'on désire effectuer les mesures par voie photographique, on utilise une chambre avec un chassis multiplicateur, l'objectif recevant la même position que précédemment; on utilise alors les méthodes habituelles de la photométrie photographique, sur lesquelles nous n'insisterons pas ici.

On pourra s'etonner de ce que nous ayons adopté la photométrie visuelle et non les méthodes photoélectriques; la raison est que les mesures visuelles ont une précision bien suffisante pour les besoins de la pratique, et que d'autre part elles sont beaucoup plus rapides et sûres. En particulier, le réglage des dispositifs photoélectriques demande un très grand soin et son maniement exige un spécialiste.

Signalons pour terminer que les mesures visuelles de la lumière parasite produite par les lentilles doivent être faites en prenant une pupille d'entrée du photomètre de diamètre inférieur à celui de l'anneau oculaire de l'instrument; il n'y a alors pas de difficultés. Il n'en est pas de même lorsqu'on fait intervenir la lumière parasite provenant des montures; la pupille d'entrée du photomètre est alors obligatoirement plus grande que l'anneau oculaire ou que la pupille de sortie, de sorte que le diamètre de la pupille de sortie du photomètre n'est pas le même selon que l'on opère sur l'objet ou sur l'image. Afin d'éviter les erreurs très importantes provenant de l'effet Stiles-Crawford, il faut que le diamètre de l'anneau oculaire du photomètre soit plus petit que 1 mm, ce qui s'obtient en donnant à ce dernier un grossissement égal au diamètre de sa pupille d'entrée, exprimé en millimètres.

MESURE DES LIMITES DE RESOLUTION ET DE L'EFFICACITE

Dans le dispositif précédent, on remplace la plage noire par le test objet; l'observateur regarde directement le champ objet de l'instrument, sans intervention d'aucun instrument grossissant auxiliaire. Lorsqu'on doit faire intervenir simultanément un champ de luminance uniforme et une source brillante, on ne peut évidemment plus séparer les mesures, comme dans le cas de la photométrie. Les deux causes doivent intervenir ensemble, et leurs caractéristiques photométriques relatives doivent être dans le rapport voulu. C'est ce qui peut nécessiter l'usage d'un arc électrique.

REALISATION D'UN TEST DONT LA GRANDEUR ET LE CONTRASTE VARIENT D'UNE MANIERE CONTINUE (fig. 9.)

On projette dans l'ouverture de l'écran E2 l'image I d'une mire M de contraste connu, au moyen d'un objectif O. L'objectif et la mire sont mobiles longitudinalement, les mouvements étant réalisés de manière que l'image reste formée constamment dans le plan de l'écran E2, tandis que sa grandeur varie de manière connue. La mire M est éclairée par transparence au moyen d'un écran diffusant E3, dont la luminance est réglée à l'égalité avec celles des autres écrans.

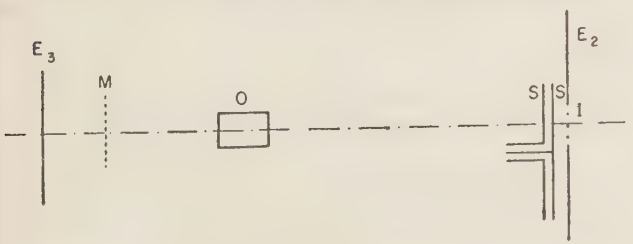


Figure 9

La variation du contraste de l'objet est obtenue en disposant au voisinage immédiat de l'image I un secteur tournant SS′ dont la face tournée vers l'instrument diffuse comme la face E2, avec la même luminance, et dont l'ouverture peut varier à volonté pendant la rotation. Ce secteur tournant est composé de deux disques concentriques juxtaposés, portant chacun 2 ouvertures opposées identiques en secteurs d'angles d'ouverture de 90°. L'ouverture résultante peut être réglée à volonté par rotation relative des 2 disques, suivant un secteur d'angle variant de 0° à 90°. Pendant la durée d'un tour, la mire n'apparait que pendant le temps de passage de l'ouverture, et pendant le reste du temps, est remplacée par une surface uniforme de même luminance que l'écran E2. Par suite de la persistance des impressions sur la rétine, la vitesse de rotation étant de 70 tours/seconde, le contraste perçu est égal à:

$$C = \alpha / 2\pi .$$

L'appareil fournit donc d'une manière continue tous les contrastes compris entre 0 et 0.5, plus le contraste 1, obtenu en observant la mire à travers l'ouverture maximum du disque, qui reste alors immobile. Il n'est pas possible d'obtenir les contrastes intermédiaires entre 0.5 et 1.

L'étude des objectifs photographiques est effectuée en plaçant une chambre sur le banc B, les différents points du champ étant étudiés par rotation de la chambre autour de l'axe A1. On compare, avec la même émulsion, l'objectif à étudier et un objectif pratiquement parfait, constitué par un Clairaut de courte longueur focale.

Les figures 2 et 3 montrent des exemples de courbes d'efficacité obtenues par cette méthode.

En ce qui concerne les instruments visuels, les appareils précédemment décrits sont utilisées pour tous les contrôles de qualité effectués par l'Institut d'Optique; on a ainsi éliminé définitivement les divergences constatées avant leur emploi entre les résultats des contrôles de laboratoire et ceux obtenus par les usagers dans l'emploi pratique des instruments.

De nombreux contrôles d'objectifs photographiques ont été effectués par la même méthode, par photographie.

REFERENCES

1. Goldberg, E.,1999, *La Formation de l'Image Photographique. Considerations Photométriques et Sensitométriques*, Publications Photographiques Paul Montel, Paris.
2. Lyot, B. et Francon, M., 1950, Etude de défauts d'homogénéité des grands disques de verre, *Rev. Opt.*, **29**, 499.
3. Francon, M. et Cagnet, M., 1950, *ibid.*, 377.
4. Goldberg, E., *loc. cit.*, p. 75 ff.
5. Dossier Brigitte, 1954, Recherches sur l'apodisation des images optiques, *Rev. Opt.*, **33**, 57, 147, 267; Errata, p. 552.
6. Saez, A., 1952, *C.R. Acad. Sci., Paris*, **235**, 148.
7. Marquet, A., 1947, Limite de résolution des émulsions et efficacité des objectifs photographiques, *Sci. Industr. photogr.*, (2ème) **18**, 129.

DEVELOPMENTS IN CATADIOPTRIC TELEPHOTO SYSTEMS

FRANK G. BACK
ZOOMAR, Glen Cove, N.Y.

ABSTRACT

Mirror optics, though known for about three centuries, have been used in general optical design only since the nineteen-thirties. An optical mirror offers numerous advantages over dioptric elements. On the other hand, it poses certain design problems which are not present in systems consisting only of lenses. By combining lenses with mirrors it is possible to produce objectives, especially telephoto systems of light weight, small physical dimensions and with an image quality not only equal but in many cases superior to images produced by standard telephoto lenses.

Mirror telescopes have been known for nearly three centuries. The well known instruments of Newton[1,2], Gregory[3,4] and Cassegrain[4,5] were all invented within one decade. After that no significant progress in the design of mirror optics was made for nearly 250 years. It was only after Bernhardt Schmidt[6,7,8] brought out his famous astrographic camera and thereby pioneered the use of lens and mirror combinations that the image forming mirror as optical component came into its own. The catadioptric systems of Maksutov[9], Bouwers[10], Slevogt[11], Sonnefeld and Baker[12], to name just a few, followed in rapid succession and today lens-mirror combinations have their definite place besides the purely dioptric systems not only as telescopes, but also as microscope and camera objectives.

The combination of lenses and spherical mirrors in a catadioptric camera objective is very advantageous if long equivalent focal length in combination with light weight and small physical dimensions are required. A lens mirror combination offers, however, considerable design problems. The aberrational contributions of a mirror, though smaller than those of an equivalent lens, cannot for obvious reasons be changed by bending. Though the extra-axial aberrations of a mirror are highly susceptible to a shift in the aperture stop, this latter is more or less indeterminate due to the blocked out centre of the mirror aperture, especially when a more extended field is to be covered. Shape and position of the dioptric components therefore have to be fitted to the optical properties of the mirrors.

These properties of a mirror can best be described by its third order aberration contributions. These third order terms give a pretty accurate picture of the mirror performance even at apertures and field angles where in an ordinary lens the higher order defects make these figures more or less meaningless as a measure of correction. The aberration contributions are not expressed as Seidel terms, or Berek coefficients, but in Conrady expressions which give the longitudinal and lateral deviations from the ideal image point. For reasons of simplicity the required aperture ray is traced

Received July 22, 1956.

from infinity and the field-or principal ray through the centre of the mirror. In a mirror the index of refraction $n = -1$. All data entering Conrady's formulae can be expressed in terms of the semi-aperture Y, the focal length F and the image height H. Conrady's aberration terms thereby become

Spherical aberration	$+ Y^2/8F$
Offence against the sine condition	$- (Y/2F)^2 H$
Saggital curvature	$+ H^2/2F$
Petzval curvature	$- H^2/2F$
Distortion	$- H^3/F^2 + H^3/F^2.$

This shows that in a mirror spherical aberration and astigmatism have the same sign as the power of the mirror, while the sine defect and the Petzval curve have the opposite sign. Distortion is, of course, zero due to the central passage of the principal ray. The two chromatic terms do not enter the picture for obvious reasons. For central passage of the principal ray the sagittal field is flat, while the tangential field and also the astigmatism measured longitudinally along the axis becomes $+ H^2/F$.

If the sine defect is presented according to Von Rohr as the difference between the theoretical focal length and the actual focal length at that particular zone, it can be shown that this difference is numerically equal to the longitudinal spherical aberration, but has the opposite sign.

The formula for the offence against the sine condition "OSC" is:

$$\text{OSC} = 1 - [Yu'(l' - l'_{pr})] / [Y \sin U'(L' - l'_{pr})],$$

where u is the paraxial aperture angle, U' the real aperture angle at the zone Y, l' the paraxial distance from the vertex of the last surface to the image point, L' the real vertex—image distance for zone Y, and l'_{pr} the paraxial distance between the exit pupil and the last lens vertex. If the aperture stop is placed in the centre of the mirror curvature, OSC becomes zero. In this case l'_{pr} equals $-2F$. (The centre of a mirror corresponds to the anti-principal points of a lens. The aperture stop is therefore imaged upon itself by the mirror.) The image distance $l' = -F$ while L' becomes $-F + LA$, the spherical aberration. Y/u' can be rewritten as F and $Y/\sin U'$ can be replaced by F_z, the real focal length at the zone Y. The condition for OSC to be equal to zero thus becomes

$$\text{OSC} = 1 - F_z F / F(F + LA) = 0$$

$$F_z = F + LA,$$

therefore real zonal focal length is equal to the nominal focal length plus spherical aberration measured longitudinally.

The foregoing shows clearly the lines to be followed in the design of a Catadioptric Telephoto Lens. If the focal length is very long as compared to the field size to be covered, a first surface primary and a first surface secondary mirror can be employed.

The higher order spherical aberrations, necessary to correct the inherent primary aberration of the mirrors, is introduced by a nearly powerless corrector plate. Usually there is not much correction required, because the spherical aberration of the primary mirror is partly offset by the opposite spherical aberration of the secondary. A dioptric, negative, rear element serves mainly to correct the offence against the sine condition. Due to the small extent of the field, fulfilment of the sine condition also means freedom from coma and its higher order variations. For the same reason, namely the small extent of the field as compared to the focal length, the decrease of astigmatism produced by the dioptric rear element is sufficient to give satisfactory resolution even in the corners in spite of the distinctly negative Petzval sum. The chromatic correction is no problem at all, because neither the powerless corrector nor the two mirrors have any chromatic defects. The relatively compact dioptric rear system can be easily achromatized within itself, and the remaining secondary colour of the rear component is not noticeable if the glass is correctly chosen.

For a more extended field, matters become more complicated. The rigid aberrational contributions of the two mirrors become less and less manageable the more the field angle increases. The inherent astigmatism and field curvature of the system become distinctly noticeable, furthermore aplanatic correction alone is not sufficient to insure freedom from objectionable higher order coma. The degrees of freedom offered by the rear system are not sufficient to take care of all these defects, and its influence on some of these aberrations is so small that even with extreme shapes it would be impossible to produce satisfactory image quality. It therefore becomes necessary to use a Mangin type mirror as primary. Such a mirror has been introduced not only because it is free from spherical aberration, but also because it can be bent like a lens without changing its power and the position of its principal points. A Mangin type mirror, though, not only introduces a slight amount of colour, due to its refracting front surface, but it also increases the negative Petzval sum of the mirror by twice the contributions of its front surface. The corrector plate therefore has to have a finite positive power and must also introduce some colour of its own to offset these aberrations.

For an even more extended field a positive rear system has to be used. In this case the corrector plate and the two mirrors have to be corrected for spherical aberration and coma and, of course, for colour. The positive rear element corrects astigmatism and flattens the field. Due to its close position to the image plane, it is sometimes possible to use a single lens because its contributions to spherical aberration and colour are negligible.

In conclusion it can be said that the use of catadioptric camera objectives is steadily increasing. This is due not only to their light weight and small physical dimensions as compared to an equivalent standard lens, but also to the fact that the image quality obtainable with a properly designed catadioptric objective is superior to that of a purely dioptric system.

I should like to thank Mr. H. Lowen for his valuable collaboration on this paper.

REFERENCES

1. Newton, Sir Isaac, 1703, *Opticks*, Proposition VIII, Problem II.
2. Chretien, H., 1927, Le télescope de Newton et le télescope aplanétique, *Rev. Opt.*, **49**, 113.
3. Gregory, J., 1663, *Optica Promota*, London.
4. Jones, R. T., 1954, Coma of modified Gregorian and Cassegrainian mirror systems, *J. opt. Soc. Amer.*, **44**, 630—633.
5. Yoder, P. R. et al., 1953, Analysis of Cassegrainian-type telescopic systems, *J. opt. Soc. Amer.* **43**, 1200—1204.
6. Schmidt, B, 1932, *Mitt. hamburg. Sternw.*, **7**, 15.
7. Caratheodory, C., 1940, Elementare Theorie des Spiegelteleskops von B. Schmidt, *Hamb. math. Einzelschr.*, No. 28.
8. Kuehn, W., 1954, Ein komafreies Spiegelsystem mit geebnetem stigmatischem Bildfeld von 5° und kurzer Baulaenge, das dem Schmidt-Spiegel mit Ebnungslinse praktisch gleichwertig ist, *Jenaer Jahrbuch*, 352—362.
9. Maksutov, D. D., 1944, New catadioptric meniscus systems, *J. opt. Soc. Amer.*, **34**, 270—284.
10. Bouwers, A., 1946, *Achievements in Optics*, Elsevier Publishing Co., 1—45.
11. Slevogt, H., 1942, Ueber eine Gruppe von aplanatischen Spiegelsystemen, *Z. InstrumKde*, **62**, 312—327.
12. Baker, J. G., 1940, A family of flat field cameras, equivalent in performance to the Schmidt camera, *Proc. Amer. phil. Soc.*, **82**, 339—349.
13. Fluegge, J., 1955, *Das photographische Objektiv*, Springer Verlag, 191—199.
14. Hekker, F., 1947, *On concentric optical systems*, Delftsche Uitgevers Mj.
15. Koehler, H., 1949, Die Entwicklung der aplanatischen Spiegelsysteme, *Astr. Nachr.*, **278**, 1—23.
16. Linfoot, E. H., 1955, *Recent Advances in Optics*, Oxford University Press, 176—228.
17. Raitiere, L. P., 1952, Computation of spherical-mirror systems by means of finite-difference equations, *J. opt. Soc. Amer.*, **42**, 960—965.
18. Schwarzschild, K., 1905, Untersuchungen zur geometrischen Optik II, Theorie der Spiegelteleskope, *Astr. Mitt. Goettingen*, **10.**

PATENTS

United States of America

Back, F. G., 1951, 2,561,774; 1952, 2,610,547; 1955, 2,701,983; 1956, 2,730,926.
Bouwers, A., 1946, 2,409,186; 1947, 2,420,348; 1950, 2,503,319; 1950, 2,504,383.
Grey, D. S., 1950, 2,520,633, 2,520,634, 2,520,635, 2,520,636.
Hayward, R., 1946, 2,403,660.
Kaprelian, E., 1945, 2,378,301; 1954, 2,685,820.
Mandler, 1955, 2,726,574.

Great Britain

Newton, G. C., 1941, 538,622.
Gabor, D., 1940, 544,694.

The Netherlands

Bouwers, A., 1941, 102,016; 1942, 104,938.

France

Bouwers, A., 1943, 883,937.
Philips, N. V., 1944, 398,315.
Paul, M. J. D., 1950, 969,797.

Germany

Mandler, 1951, 824,558.

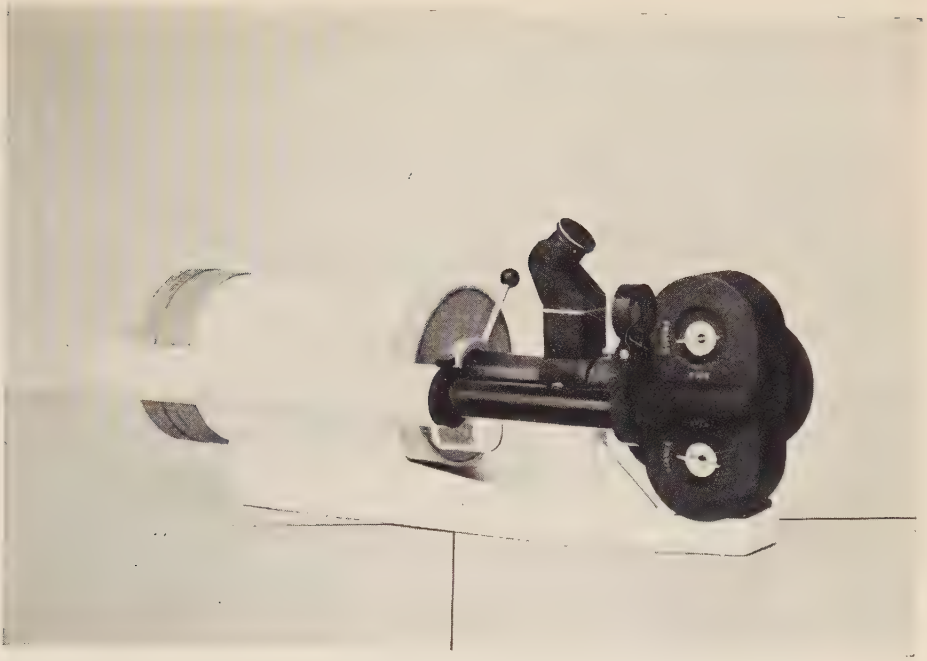


Figure 1
Catadioptric telephoto lens as used in U.S. military installations.

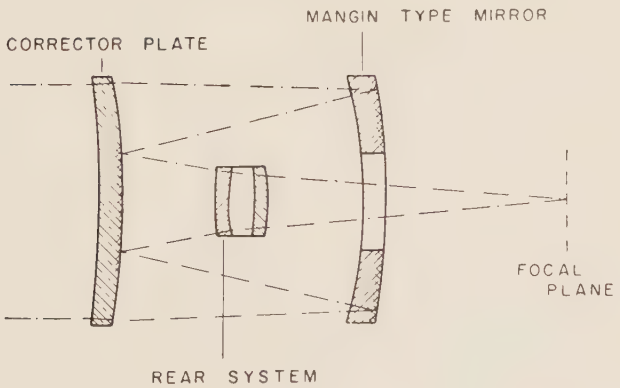


Figure 2
Schematic of a typical catadioptric telephoto lens.

REFRACTIVE INDICES AND OPTICAL DISPERSION OF GASES IN THE INFRARED REGION

J. H. Jaffe

The Weizmann Institute of Science, Rehovot

ABSTRACT

Optical dispersion is intimately related to optical absorption and both yield similar information about the structure of matter. The work which has been done on the measurement and interpretation of dispersion of gases in the infrared region is reviewed.

It is explained that important new results are to be obtained from improved infrared dispersion techniques. By linking together an infrared refractometer of very high sensitivity and a monochromator of the highest resolving power it seems possible to obtain data that are at present beyond the reach of conventional absorption spectrometry. Current efforts in this direction are described. A preliminary account is given of the large infrared refractometer-spectrometer now under construction in Rehovot.

INTRODUCTION

The purpose of this paper is to survey briefly the work that has been done on the measurement of the refractive indices and dispersion of gases in the infrared region. Optical dispersion is of course intimately related to optical absorption. Sharp 'wrinkles' in the dispersion curve are to be found at the same wavelengths as strong absorption bands, and both the wrinkles and the bands yield similar information about the structure of matter. However, owing to limitations of instrumentation the one type of measurement is often more useful or accurate than the other. Broadly speaking, it may be said that the absorption spectrum is naturally suited for the measurements of the absorption frequencies, whereas the dispersion curve very often is superior for determinations of quantities associated with absorption intensities.

EARLY MEASUREMENTS

Koch[1], in 1909, reported measurements of the refractive indices of methane and carbon dioxide at 6.71 and 8.68 μ, using a Jamin interference refractometer with rocksalt optics. A complete dispersion curve of carbon dioxide in the region 1 to 13 μ was given by Statescu[2]. His method involved measuring the refraction of the sample placed in a hollow prism made up of a pair of rocksalt plates. The dispersion curve has a strong anomaly in the region of the absorption band at 4.3 μ and the abnormally low values at 13 μ — the wavelength limit of the measurements — are

Received July 30, 1956.

indicative of a second strong anomaly at 15 μ. Further work on CO_2 with a similar hollow prism refractometer was carried out by Wetterblad[3].

The above workers confined themselves to experimental measurements and made no attempt at a theoretical interpretation of their results.

VAN VLECK'S ANALYSIS

Great stimulus to work on infrared dispersion of gases was given by Van Vleck[4] by an analysis of discrepancies in the measurements of the temperature invariant parts of the dielectric constants of CO_2 and HCl.

The static dielectric constant ε of a gas may in general be expressed in the form:

$$\varepsilon = 1 + 4\pi N(\alpha + \mu^2/3\mathrm{kT}) . \tag{1}$$

N is the number of molecules per cc; μ is the permanent dipole moment and the term $4\pi N\mu^2/3kT$ represents the contribution to ε due to the orientation of the permanent electric dipoles in a field. This is the temperature dependent term; the tendency of the dipoles to line up in the direction of an impressed field becomes less with increasing temperature, because of the competitive influence of thermal agitation. The quantity α is the polarizability and the term $4\pi N\alpha$ is mainly the contribution from distortion of the distribution of electronic charge around the molecules. Except for a density correction to N this term is independent of temperature. Since the electrons are of small mass, capable of following very high frequency oscillations of an impressed field, and the permanent dipoles on the other hand cannot do so, it should be possible to obtain the distortion polarization by itself by measuring the dielectric constant in such a high frequency field — at optical frequencies, in fact. That is to say, at optical frequencies the value of the dielectric constant should be equal to

$$1 + 4\pi N\alpha .$$

The Maxwell relation[5] equates the dielectric constant to the square of the refractive index n

$$n^2 - 1 = 4\pi N\alpha . \tag{2}$$

However, the refractive index n depends to some extent on frequency — that is, most materials display optical dispersion — so that the precise frequency must be specified in asserting the equality (2).

Since the measurement of dielectric constants can only be made with static or quasi-static fields, it is necessary to extrapolate the optical dispersion curve to zero frequency in order to obtain the value of the index n_∞ for computation of the temperature independent part of ε. Now this 'refractivity method' is not the only

way in which the temperature independent term can be determined. For example, it can also be calculated from extrapolation of (1) to infinite temperature. Comparison of these two calculations for HCl revealed a considerable discrepancy.

Van Vleck cites the value for $(n_\infty^2 - 1)$ by extrapolation of optical data to zero frequency as 871×10^{-6}, whereas $4\pi N\alpha$ by extrapolating ε to infinite temperature is 1040×10^{-6}. The difference between these two values is

$$4\pi N\alpha - (n_\infty^2 - 1) = 169 \pm 12 \times 10^{-6}. \tag{3}$$

It was suggested that this discrepancy may well be due to the fact that the optical dispersion curve for the visible region was wrongly extrapolated to zero frequency and in fact dispersion anomalies in the infrared region corresponding to atomic polarizations should be taken into account. This correction was expressed in the form

$$\Sigma \frac{Ne_i^2}{3\pi\mu_i\nu_i^2} \tag{4}$$

where ν_i are vibrational frequencies and the summation is carried out over all such frequencies which are active. The quantities e_i are the effective charges corresponding to the vibrations ν_i and μ_i are reduced masses. Now the ν_i in expression (4) can be determined and μ_i are directly calculated. Values for e_i can be computed from the absolute intensities of absorption, and indeed for HCl the experimental data of Bourgin[6] yield the value 0.95×10^{-10} e.s.u.[7] Substituting this value in (4) one finds that the difference $4\pi N\alpha - (n_\infty^2 - 1)$ is equal to 1.5×10^{-6}. That is to say, the effective charge computed from the absorption data barely accounts for one percent of the discrepancy of 169×10^{-6} between the optical refractive index extrapolated to zero frequency and the temperature invariant part of the dielectric constant. It was felt that this difficulty may have arisen owing to the supposedly unreliable measurements of absolute integrated absorption of the HCl bands, that is, it was thought that the effective charge computed from the absorption curve may have been very much too small.

ROLLEFSON'S WORK

At the direct instigation of Van Vleck an attempt was made by Rollefson[8] to determine the value of the effective charge by means of the infrared dispersion curve.

Rollefson's refractometer was similar to that of Statescu mentioned above. The light from a monochromator was passed through a hollow prism made up of a pair of rocksalt plates set at 90° to each other. The gas under examination was put into this prism and the deviation of the light beam was measured in a simple direct way. The accuracy of index determination was about 5×10^{-6}.

The data were fitted to a Kramers—Heisenberg dispersion formula

$$n^2-1 = \frac{8\pi N}{\Sigma \exp(-W_m/kT)} \Sigma \frac{|M_{mn}|^2 \nu_m \exp(-W_m/kT)}{h(\nu_m^2-\nu^2)} \tag{5}$$

M_{mn} is the transition moment from mth to nth state and ν_m is the corresponding frequency. W_m is the energy of the mth state.

Despite the fact that the wavelength resolution of his equipment was quite poor, Rollefson recognized that it was necessary to include in the dispersion formula contributions corresponding to the frequencies of the individual vibrational-rotational lines which make up the fundamental band at 3.5 μ. It was further necessary to take into account their relative intensities.

In its final form the refractive index at a given frequency was expressed as the sum of three terms

$$n^2-1=(n^2-1)_{C\&C}+f(\nu)e_{eff}^2+g(\nu)\mu^2. \tag{6}$$

The suffix C&C means that the first term is the contribution of the electronic polarization and is computed by extrapolation of the dispersion curve of Cuthbertson and Cuthbertson for the visible region[9]. The second term is proportional to the square of the effective charge and the third is proportional to the square of the permanent dipole moment μ. The significance of this last term is that molecular rotation is also taken into account. The values of e_{eff} and μ are obtained from the experimental data by a process of successive approximation, and the final values given are

$$e_{eff}=1.00\times10^{-10} \text{ e.s.u.}$$

$$\mu \ =1.18\times10^{-18} \text{ e.s.u.}$$

There is a striking agreement between the value of the effective charge and that deduced from Bourgin's absorption data (0.95×10^{-10} e.s.u.). This dispersion work did not, then, show that the absorption measurements were so very unreliable as was expected. It is a historical fact that the immediate objective which was in mind at the initiation of modern infrared dispersion measurements, turned out to be imaginary. The value of the effective charge computed from integrated absolute absorption data was apparently excellent. The great discrepancy (3) seems to have been due to an error in the then accepted value of the dipole moment of HCl which was that of Zahn[10]: $\mu = 1.034 \times 10^{-18}$ e.s.u. Rollefson's own value is in any case too high. The value accepted today[11] is in the region of 1.08×10^{-18} e.s.u. and this well accounts for the missing 169×10^{-6} of equation (3).

It is of interest to mention the case of CO_2. Here again Van Vleck discussed a large discrepancy between the measured and expected values of the effective charges for the fundamental vibrations. This has since been ironed out. The case is unusual because the infrared dispersion data were available[2,3] before the absolute absorption intensities. A dispersion formula for CO_2 was worked out by Fuchs[12] which took into account the infrared part of the dispersion curve as given experimentally by Koch and Statescu. This formula called for effective charges of 2.28 and 0.61 times the electronic charge for the 4.3 μ and 14.9 μ vibrations respectively. The values of the effective charges derived from absorption data[13] were less than these by a factor of about five. However, improved measurements of absorption by Martin and Barker[14] and more recently by Thorndike[15] agree in estimation of the effective charges very closely with those of the dispersion formula of Fuchs. This is a clear cut example of the superiority of the dispersion curve over the absorption spectrum in deducing data associated with intensities. The old and primitive dispersion measurements on CO_2 have been found to be at least as good as modern intensity measurements.

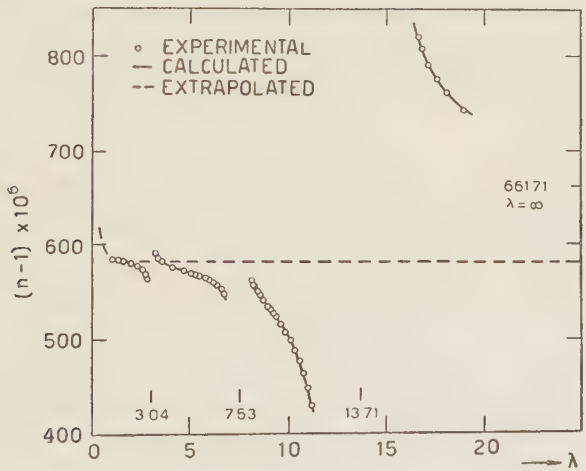


Figure 1

Typical dispersion curve in the infrared region. Acetylene at N.T.P. (From Kelly, Rollefson and Schurin[18]).

In addition to the work on HCl Rollefson and his collaborators have used the same refractometer to furnish dispersion data on methane[16], deuterium chloride, nitrous oxide, ethylene[17], and acetylene[18]. In the case of acetylene the thread of the discussion is as follows.

The refractive index n at a frequency ν was expressed as a series of terms:

$$n-1 = \Sigma \frac{C_i}{(\nu_i^2 - \nu^2)} \tag{7}$$

where the C_i and ν_i are constants. For the visible region Watson and Ramaswamy expressed the index at N.T.P. by a single term essentially as follows:

$$(n-1)_{R\&W} = 5469 \times 10^3/(9403 \times 10^6 - \nu^2), \tag{8}$$

where the frequency ν is in cm^{-1}. Rollefson found that the infrared refractive indices over the 1 to 20 μ region could be fitted to the formula

$$n-1 = (n-1)_{R\&W} + \Sigma \frac{A_i}{2\pi^2 c(\nu_i^2 - \nu^2)} \tag{9}$$

where the first term is (8), the extrapolated values of Watson and Ramaswamy's formula. In the succeeding terms, A_i are the integrated absorption coefficients of the bands situated at frequencies ν_i. Fundamental, overtone and combination bands were taken into account. Although it was not found necessary to include terms due to each individual rotational line as in the case of HCl, a distinction was made however between the P, Q and R branches of a band. For example, the ν_5 band was represented by three terms associated with three frequencies ν_5', ν_5'' and ν_5''' corresponding to the P, Q and R branches respectively. The infrared contributions to the value of the refractive index extrapolated to zero frequency amount to 80.1×10^{-6}. That is, from (9),

$$n - 1 = 661.7 \times 10^{-6},$$

which corresponds to a dielectric constant

$$\varepsilon - 1 = 1323.9 \times 10^{-6}.$$

This value agrees with the experimentally determined static dielectric constant of acetylene[20] (1.001330 at N.T.P.) to within the limits of experimental error.

MODERN DEVELOPMENTS

The work summarized in the above paragraphs was almost exclusively designed to determine optical dispersion curves as a whole over a large portion of the infrared region. As a rule the aim of these determinations was to compute the contributions of molecular vibrations and rotations to dielectric constants. In many cases values for the effective charges associated with the individual absorption bands were also deduced. However, it is hard to point to anything essentially new in these dispersion measurements that has not been, or could not have been deduced from absolute integrated absorptions. The state of affairs was summarized by Rollefson in his last paper[18]: "For simple molecules whose absorption is strong enough, measurements of infrared dispersion provide a sensitive and accurate method of determining the intensities of infrared bands."

Although there are great experimental difficulties in making accurate measurements of absolute absorptions, there has been much improvement in recent years in the measurement of the *total integrated absorption* of a band as a whole. In the well established method of Wilson and Wells[21] the band under examination is broadened by pressurizing it with a non-absorbing gas. The intensity is obtained by measuring the integrated apparent absorption for unit pressure at a series of partial pressures. Essentially this procedure makes it possible to apply the instrumental slit corrections. Roughly speaking, the absorption band is pressure broadened so that it is made broad compared with the spectral slit width of the instrument.

The opposite procedure cannot be adopted. If the band is left narrow while the spectral slit width is reduced by improving the resolving power of the spectrometer, then the situation is actually aggravated because of the resolution of the rotational fine structure. And no spectrometer has a spectral slit width which is negligible by comparison with a single rotational line.

Much interesting information is to be derived, however, from the intensities of individual rotational lines, and it is believed that herein lies an opportunity to apply improved dispersion techniques. By linking together an infrared refractometer of very high sensitivity and a monochromator of the highest resolving power, it should be possible to measure accurately what is equivalent to the absolute integrated absorptions of individual rotational lines. Such data are at present beyond the reach of conventional absorption methods.

The need for improving dispersion measurements and particularly linking them with increased wavelength resolution was appreciated independently both here in Rehovot and by F. Legay in Paris[22].

Legay's refractometer is essentially a Twyman-Green interferometer with rocksalt lenses and a lithium fluoride beam splitter coated with a semi-reflecting layer of selenium. The interferometer arms accommodate gas cells of 25 cm in length. Measurements of refractive index can be made with an accuracy of 5×10^{-7} at 3.4 μ or about ten times that claimed by Rollefson. The refractometer is used in conjunction with an excellent monochromator of the Pfund mirror type, incorporating a plane grating ruled with 300 lines/mm. The detector is a lead selenide cell, cooled with liquid air. It is claimed that the instrument resolves 0.3 cm^{-1} at 3.5 μ.

There are two alternative methods of using this type of refractometer. In the first method the monochromator is set at a given wavelength. Initially the cell in one of the arms of the interferometer is evacuated and the sample gas is allowed to flow steadily into it. Interference fringes are counted as this process continues. From the exact number of fringes which correspond to the final pressure and temperature it is possible to compute the refractive index directly.

The second method is better adapted for the determination of the general form of the dispersion curve, but it is probably less accurate and the record is harder to interpret. The gas pressure in the cell is kept constant while the monochromator

is swept through a region of the spectrum. An example of the resulting channelled spectrum of a section of the fundamental band of HCl at 3.5 μ is shown in Figure 2.

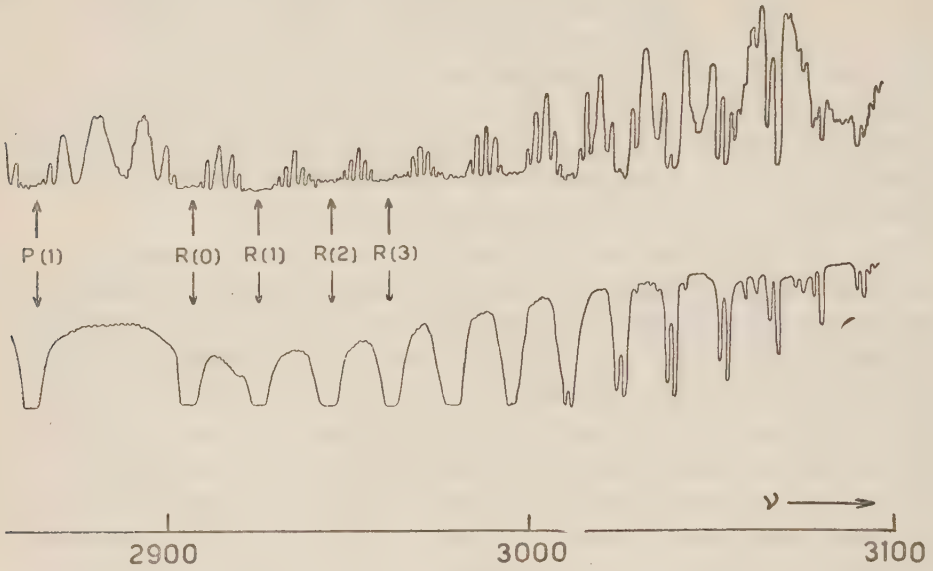


Figure 2

Portion of a channelled spectrum produced with a two beam interferometer linked to a monochromator of high resolution. HCl gas at 582 mm pressure and 27.8°C. (From F. Legay[29]).

HCl was a good molecule to use for preliminary investigations because of the wide spacing of the rotational lines in the vibration-rotation band. In a first attempt at accurate measurements, Legay determined the dispersion in the spectral region between P(1) and R(0) of the fundamental band. His dispersion curve is reproduced in Figure 3.

Legay fitted his data to a Kramers-Heisenberg dispersion formula involving transition moments of a harmonic oscillator and rigid rotator. In a later paper[23], however, he interpreted the same data together with additional data taken on the harmonic band at 1.7 μ, in terms of a refined theory of Herman and Wallis[24] which embodies the most up-to-date treatment of vibration-rotation interactions. Accordingly, he was able to write down in explicit form the dipole moment $\mu(r)$ of the HCl molecule as a function of the intermolecular distance. The expression contains a quadratic term.

$$\mu(r) = 1.092 + 0.953(r - r_e) - 0.275(r - r_e)^2, \tag{10}$$

where $(r - r_e)$ is in A and $\mu(r)$ is expressed in Debye units.

Superiority of dispersion measurements over those of integrated absorption intensities is due in part to the fact that the patterns of dispersion curves are much more extensive than those of the corresponding absorption lines. Spectral slit width

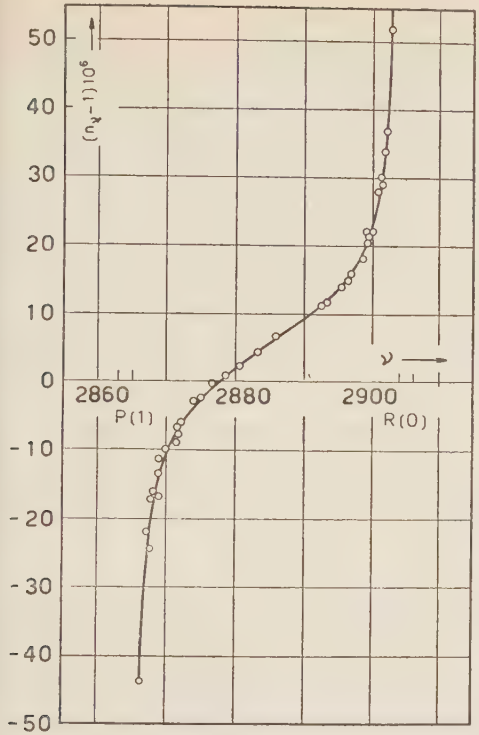


Figure 3
Strong dispersion of HCl gas at 760 mm pressure and 15°C between individual lines of the fundamental band. (From F. Legay[29]).

corrections are therefore easier to apply. Furthermore, it seems probable that the all important correction for Doppler broadening can be made more accurately in the case of dispersion, but this is a question which has not yet been fully investigated. In addition the problem of background correction is simplified. The magnitude of a given wrinkle in the dispersion curve can be determined with accuracy without detailed knowledge of the general 'background of refraction'. This point is illustrated in Figure 4. An absorption line is drawn in Figure 4(a) which has a measured height A. There is no direct and accurate way of knowing what fraction of A is actually due to the background S. In Figure 4(b) we depict the corresponding dispersion curve D. The quantity on the dispersion curve which corresponds approximately to the height of the absorption line is the difference in refractive index between its minimum and maximum. Should the dispersion curve be superimposed upon an extraneous background (e.g. curve D′), then this does not substantially influence the determination of the height $n_{max} - n_{min}$.

Figure 4(b) serves to illustrate a further point. The slope of the dispersion curve on either side of the actual wrinkle is shown to be quite steep due to the influence, say, of neighbouring bands. This does not interfere with the computation of the contribution of the wrinkle itself, provided that the slope is not so steep or the wrinkle so weak as to be obliterated completely. Herzfeld[25] has shown how to interpret such an element of a dispersion curve superimposed upon a steep back-

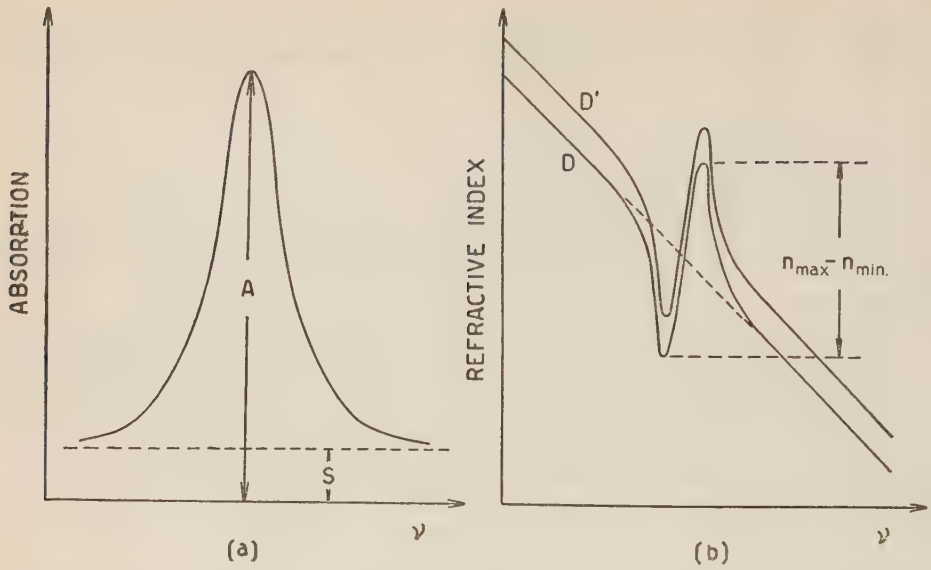


Figure 4
Influence of extraneous background on (a) absorption line and (b) dispersion 'wrinkle'.

ground. His treatment assumes that the general gradient (dotted in Figure 5(b)) is the constant over a small frequency interval of order of magnitude of the absorption line half width.

THE REHOVOT REFRACTOMETER-SPECTROMETER

A large infrared gas refractometer is at present under construction in Rehovot. It is of the hollow prism type — essentially the same as that used by Statescu[2], Wetterblad[3] and Rollefson[8]. The choice of this design was made after careful consideration of a number of possible alternatives.

What is the most sensitive type of gas refractometer? Ingelstam[26] has described a phase contrast refractometer suitable for work with gases which gives differential indices with an accuracy of one part in 10^8 or better. It was not felt that this instrument could be successfully adapted for use in the infrared region. The method involves a drastic division of wavefront which would be quite unsuccessful in a spectral region where the level of radiant energy is low. A further disadvantage of the Ingelstam refractometer is that it involves a wavelength sensitive element in the form of a phase plate.

Of interferometric refractometers there is the choice between two main types — the multiple beam Fabry-Perot etalon with the sample between the plates, or the two beam Twyman-Green interferometer. Advantages of the Fabry-Perot etalon are those of compactness and relative simplicity in construction. The main disadvantage is one which is only apparent to those who have had direct experience

with the use of this interferometer as a refractometer: it is very difficult to eliminate the thermal disturbances which arise during the pumping out or the filling of the etalon. Rank[27] has alluded to this problem in connection with his work on the refractive index of air. In the infrared, the semireflecting films with which the Fabry-Perot plates must be coated are either of poor reflectivity or strong absorbers, or — if made of multiple dielectric layers — the reflectivity and also the phase change on reflection are wavelength dependent. Difficulties with the semireflecting films constitute a second reason against using the Fabry-Perot etalon. In contrast the Twyman-Green interferometer is relatively free from troubles. A semireflecting film is of course needed in this type of instrument as a beam splitter, but the reflectivity required is only 50% and satisfactory films which meet this requirement are readily made. It is believed that these are the considerations which led Legay to choose a Twyman-Green type of interferometer.

Let us now examine the straightforward class of prism refractometers in which an angle of refraction is measured directly. It appears that the simplest form of instrument of this type is in fact the hollow prism refractometer. A possible alternative is a solid prism immersed in the gas sample, but there is no advantage in it. The pair of windows needed for the gas enclosure may just as well be used to build a hollow prism. Critical angle refractometers are also feasible for gases and such an instrument could probably be designed for the near infrared. However, the construction would be costly and in any event calculation shows that the sensitivity could not be expected to exceed one part in 10^5.

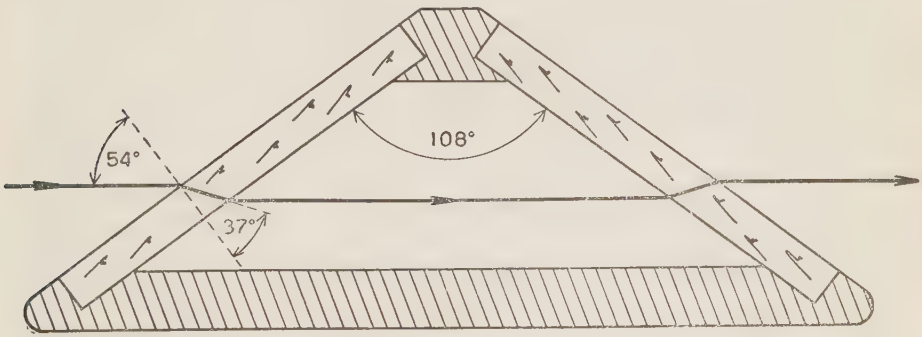


Figure 5

Section through a hollow prism. The actual deviation is very small. In the position of minimum deviation the beam traverses the prism parallel to its base.

A hollow prism made up of a pair of optical windows is sketched in Figure 5. The best disposition for the prism in the optical system is that of minimum deviation, since in it the prism has a maximum aperture, a maximum light transmission and causes minimum astigmatism. The position of minimum deviation for a gas prism, as for a solid prism, is the symmetrical position.

The prism in the Rehovot refractometer is set at minimum deviation in a five metre Littrow system, and the beam can be multiply passed[28] through it up to four times.

Consider a refraction at a face of the prism. The plane parallel window exerts no influence and the refraction may be regarded as taking place directly at an air-sample interface. Let the angle of incidence be ϑ and let the index of the sample gas be n. A change in n will change the angle of refraction by an amount

$$(\Delta n/n)\tan\vartheta.$$

At each pass through the Littrow system the beam enters the prism, emerges and then retraces its path — 4 refractions. After four passes there is a total of 16 refractions, so that the combined deflection $\Delta\vartheta$ is

$$\Delta\vartheta = 16(\Delta n/n)\tan\vartheta. \tag{11}$$

If $\Delta\vartheta$ is set equal to the smallest detectable angular interval, the equation (11) gives the smallest measurable index change Δn. It is difficult to predict $\Delta\vartheta$ from theoretical considerations alone because much depends on the quality and stability of the optical system as a whole. A skilled observer, working in the visible region with a first class spectrometer, is able to make a symmetrical setting on a suitable spectral line with an accuracy forty times better than the corresponding Rayleigh resolving power. It is believed that a conservative estimate for $\Delta\vartheta$ in our case is

$$\Delta\vartheta = \lambda/10D\cos\vartheta, \tag{12}$$

where λ is the wavelength of the light used and D is the diameter of the prism windows. Putting $\Delta\vartheta$ of (12) in (11) we obtain

$$\Delta n = \lambda/160D\sin\vartheta. \tag{13}$$

With regard to the choice of prism angle, equation (13) indicates that it should be as large as possible. However, the larger the angle of incidence, the larger are the Fresnel reflection losses. These reflection losses increase gradually with angle of incidence up to the polarizing angle, but beyond it they increase rapidly. It is therefore a sensible compromise and in fact standard practice to set the angle of incidence equal to the polarizing angle. The hollow prism of the Rehovot instrument can be fitted with a pair of borosilicate crown glass plates 200 mm in diameter (intended for work in the 1.5 μ region), or alternatively a pair of lithium fluoride plates 140 mm in diameter (intended for the 3 μ region). The polarizing angle for lithium fluoride ($n = 1.37$) is 54° and the prism angle is accordingly 108°.

It is interesting to compare this sensitivity (equation 13) with that of a two-beam interferometer. This is simply

$$\Delta n = (\lambda/2t)\Delta m, \tag{14}$$

where t is the length of the gas cell in one of the arms of the interferometer and Δm is the fraction of a fringe which is just detectable. In Legay's instrument t = 25cm. Let us take 0.05 as a probable value for Δm. Table I shows the estimated respective sensitivities of the prism and interference refractometers for two wavelength regions, and it is seen that they are substantially the same. Given equal sensitivities the prism instrument is certainly preferable. It provides data which are simple to interpret in a form well suited to direct recording. There seems to be a good prospect of recording a dispersion curve automatically as simply and quickly as it is now usual to record absorption spectra.

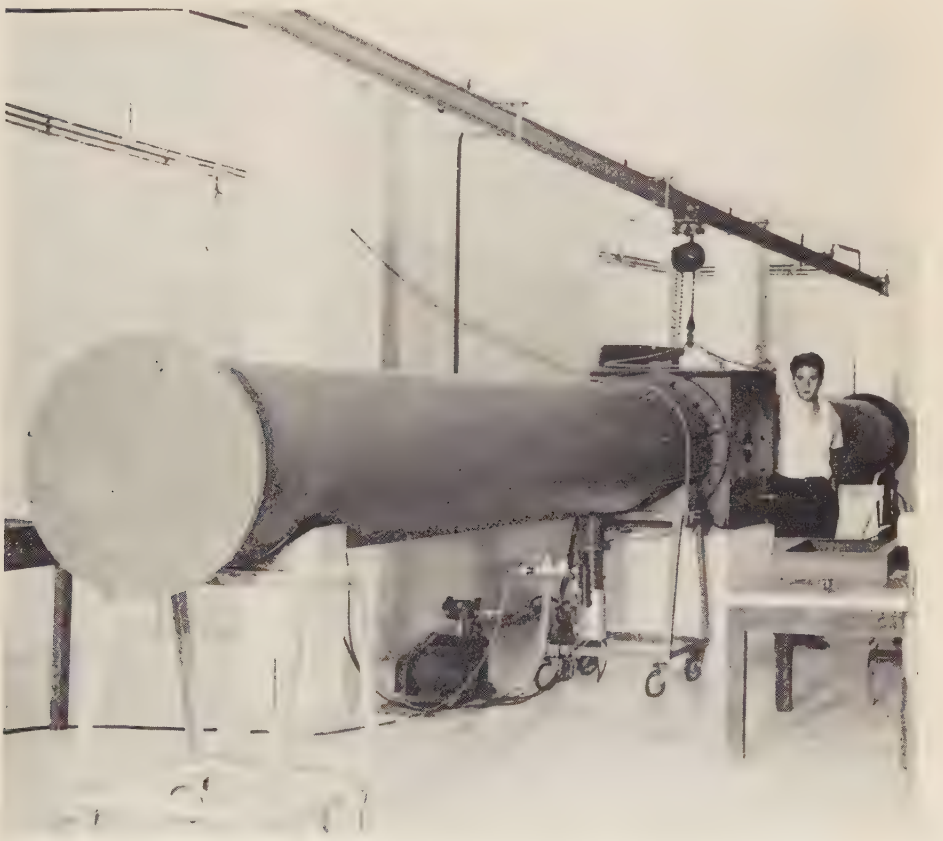


Figure 6

General view of the refractometer-spectrometer under construction in Rehovot as it stood on July 1st, 1956. The trollies are to be replaced by permanent piers. The disposition of the pumps is temporary.

TABLE I

Comparison of sensitivities of two large gas refractometers for the infrared. For 1.5 μ *the hollow prism is made of glass plates, diameter* 200 *mm. For* 3 μ *the plates are of LiF, diameter* 140 *mm.*

Refractometer	1.5 μ region	3 μ region
Prism	7.2×10^{-8}	2.3×10^{-7}
Interferometer	1.5×10^{-7}	3.0×10^{-7}

The Rehovot refractometer is being built together with a five metre grating monochromator of high quality. The entire instrument is enclosed in a casing so that it can be evacuated. Figure 6 shows a view of the instrument as it stood on July 1st, 1956. Full details of the design and construction will be published in due course.

REFERENCES

1. Koch, J., 1909, *Nova Acta Soc. Upsala*, **2**, (5).
2. Statescu, J., 1915, *Phil. Mag.*, **30**, 737.
3. Wetterblad, T., 1924, Dissertation, Upsala. Quoted in ref. 8.
4. Van Vleck, J. H., 1932, *Electric and Magnetic Susceptibilities*, Oxford.
5. Korff, S. A. and Breit, G., 1932, *Rev. mod. Phys.*, **4**, 471.
6. Bourgin, D. G., 1927, *Phys. Rev.*, **29**, 794; 1928, *ibid.*, **32**, 237.
7. Dennison, D. M., 1928, *ibid.*, **31**, 503.
8. Rollefson, R. and Rollefson, A. H., 1935, *ibid.*, **48**, 779.
9. Cuthbertson, C. and Cuthbertson, M., 1913, *Phil. Trans. roy. Soc.*, **213A.**
10. Zahn, C. T., 1924, *Phys. Rev.*, **24**, 400.
11. Le Fevre, R. J. W., 1953, *Dipole Moments*, Methuen.
12. Fuchs, O., 1927, *Z. Phys.*, **46**, 519.
13. Dennison, D. M., 1926, *Phil. Mag.*, **1**, 195.
14. Martin, P. E. and Barker, E. F., 1932, *Phys. Rev.*, **41**, 241.
15. Thorndike, A. M., 1947, *J. chem. Phys.*, **15**, 868.
16. Rollefson, R. and Havens, R., 1940, *Phys. Rev.*, **57**, 710.
17. Hammer, C. F., 1948, Thesis, Wisconsin. Quoted in ref. 18.
18. Kelly, R.L., Rollefson, R. and Schurin, B., 1951, *J. chem. Phys.*, **19**, 1595.
19. Watson, H. E. and Ramaswamy, K. L., 1936, *Proc. roy. Soc.*, **156A**, 144.
20. Watson, H. E., Rao, G.G. and Ramaswamy, K. L., 1934, *ibid.*, **143A**, 558.
21. Wilson Jr., E. B. and Wells, A. J., 1946, *J. chem. Phys.*, **14**, 578.
22. Legay, F., 1955, *C. R. Acad. Sci., Paris*, **240**, 174.
23. Legay, F., 1956, *ibid.*, **242**, 1593.
24. Herman, R. C. and Wallis, R. F., 1955, *J. chem. Phys.*, **23**, 637.
25. Herzfeld, K. F., 1939, *J. opt. Soc. Amer.*, **29**, 355.
26. Ingelstam, E., 1951, *Nature, Lond.*, **168**, 960.
27. Rank, D. H. and Shearer, J. N., 1954, *J. opt. Soc. Amer.*, **44**, 575.
28. Walsh, A., 1951, *Nature, Lond.*, **167**, 810.
29. Legay, F., 1955, *Nuovo Cim.*, **2**, Suppl. 3, 781.

PHOTOELECTRIC MEASUREMENT OF INTERFERENCE FRINGE "VISIBILITY" (THEORY AND EXPERIMENTS) AND EFFECTIVE LENGTH CORRECTIONS FOR PRECISE LENGTH MEASUREMENTS IN INTERFEROMETERS WITH CONTINUOUSLY MOVING MIRRORS

G. W. Stroke

Spectroscopy Laboratory and Research Laboratory for Electronics, Massachusetts Institute of Technology, Cambridge, Mass.

ABSTRACT

Photoelectric receptors and electronics have greatly increased the accuracy and simplicity of interferometric measurements over large path differences which extend up to several hundreds of millimeters. However, for the length measurements with precisions of 1 part in 10^8 that are now being performed it is necessary to investigate the conditions which are required to assure adequate information content and visibility of the fringe signals the amplitude of which decreases with path difference. It is shown that practical source apertures lead to a predictable error in the effective length measured in terms of a given monochromatic radiation due to flux integration over a range of angles through the interferometer determined by the aperture. Theoretical fringe visibility laws have been derived for sources that are perfectly monochromatic, Doppler broadened or Doppler broadened self-absorbed. Comparison with experimental values demonstrates that the signal-to-noise ratio of 30:1 which has been obtained at path differences as great as $\pm$ 320 mm by adequate source cooling is a result of achieving proper 5461 Å line shape by suppressing the self-absorption which is large even for a 5 mm bore Hg 198 electrodeless source at room temperatures.

INTRODUCTION

Interferometric systems in association with photoelectric detectors and electronics are being called upon to perform length measurements of a heretofore unattained precision. Dimensions of atomic diameters of the order of 1 Å are being frequently quoted in measurements with waves from 2,000 to 7,000 Å long.

Dynamical use of interferometers, that is to say use of interferometers with a moving mirror element or varying medium, has made it imperative to examine again in more detail the fringe visibility problem, some aspects of which have once fascinated Michelson[1] and Lord Rayleigh[2]. When the receptor in an interferometer is a photoelectric tube with associated electronics, the interferometer will produce an electric signal the amplitude of which varies basically in a sinusoidal manner with path difference when the mirror element is moving with uniform velocity, or

Received May 20, 1956.

Bull. Res. Counc. of Israel, Vol. 5C, 1957.

when the refractive index of the medium is varying in the corresponding manner. The AC output of the photoelectric tube is directly observable on an oscilloscope and can be used functionally for servo-mechanical control, for example[3]. The amplitude of the AC output is found to decrease with increasing mirror separation, or path difference, with the result that at great path differences no fringe signal is observed, the amplitude becoming small as compared to the detector noise level. The decrease of fringe modulation with path distance is accessible to theoretical investigation and to practical verification. These studies are needed to assure proper information content of the signals and to extend the interferometer operating range. Modern apparatus, such as the M.I.T. diffraction grating ruling engine[3] and associated photoelectronic equipment, allow very accurate visibility measurements to be made automatically and with great speed, a visibility curve measurement over path differences of 600 mm taking less than one hour. While Michelson and Lord Rayleigh were interested in interference fringe visibility perceived by the eye as a receptor, it is found that the situation has to be looked at differently when photoelectric tubes are the receptors. In fact, while the eye sees fringe maxima and minima by contrast differentiation of neighbouring fields, the phototube provides the AC fringe signal itself: however, it integrates the total, say nearly monochromatic, flux falling upon it, regardless of its contrast or brightness distribution. Figure 1 shows how a typical fringe signal of high signal-to-noise ratio appears on an oscilloscope. This particular figure corresponds to a signal produced by the green line (5461 Å) of a Hg 198 source at a path difference of 200 mm, the mirror moving with a velocity of about 2.5 mm per hour. The fine inverted trace is produced by a small generator and provides the reference phase for servo-mechanical control[3] of the moving carriage to which the interferometer mirror is attached.

When length measurements of the order of several hundred millimeters are being performed dynamically, we find that even a properly adjusted interferometer might lead to errors resulting:

(a) from an apparent change in wavelength of a perfectly monochromatic source, due to finite source aperture;

(b) from phase shifts in high order fringes (that is to say, in some of their characteristic phase, maximum or minimum for example) with this same perfectly monochromatic source;

(c) from phase shifts due to a rapid decrease of fringe modulation with a "broadened" source, the broadening being more influenced by the actual line shape than merely by an assumed Doppler form. In this respect self-absorbed lines which are found in mercury electrodeless discharge tubes of only 5 mm bore, at room temperature, are particularly dangerous.

Extensive experience at the Massachusetts Institute of Technology on interferometric control of the M.I.T. diffraction grating ruling engine[3], and the problem created by a new dynamic measurement of the velocity of light[5] to be referred to later, has

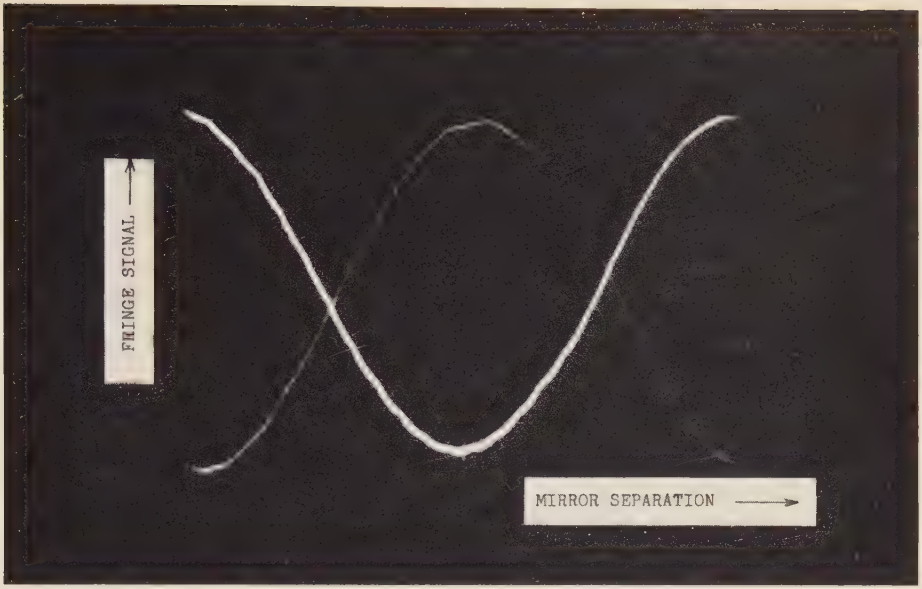


Figure 1

Oscilloscope trace of fringe signal (at 200 mm path difference). The abscissa of this figure shows the mirror separation and the ordinate the corresponding fringe signal. The fat trace corresponds to the fringe signal current from the photo-multiplier tube over a mirror separation of 1 fringe at the wavelength of 5461 Å. The inverted fine trace, with the maximum at the top of the figure, corresponds to the reference signal used in both the translation or rotation servo-control of the M.I.T. ruling engine blank carriage. The high signal to noise ratio at this path distance, appearing as a clean fringe signal, is to be noted. The signal amplitude is sinusoidal with mirror separation, or path distance, but the amplitude also decreases with increasing path distance in a manner described by equations (4) and (7). This variation of amplitude of the sinusoidal fringe signal with path distance is called a visibility curve and is shown in Figures 2 and 9 for specific cases of source line shapes.

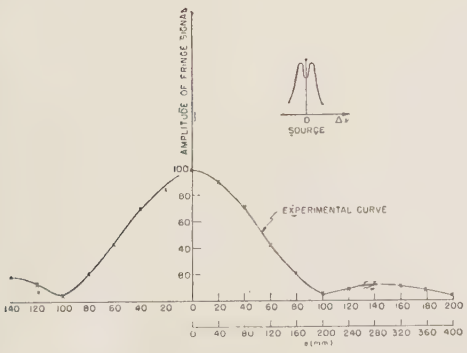


Figure 2

Experimental visibility (fringe signal amplitude) curve with self-reversed source. This curve shows the variation of the amplitude of the fringe signal with path distance. The secondary maxima at about $\pm$ 280 mm path difference e are due to the fact that the source was self-reversed as shown in Figure 8, and as sketched in the upper right hand corner here.

provided the background incentive for the calculations and experiments to be described hereafter. An early experimental visibility curve, shown in Figure 2, was obtained with a very well adjusted interferometer. There were the unexpected minima at 100 mm and secondary maxima at around 140 mm mirror separation (path differences 200 mm and 280 mm respectively) that led us to investigate the theoretical fringe modulation laws. The cause of the unexpected minima was found to reside in a self-reversal of the source operated at room temperature, as has been subsequently demonstrated. It is well known, and we have successfully verified, that great improvements in fringe modulation, or visibility, can be obtained by proper cooling of the source. As a result use of interferometers at large path differences becomes possible. One should note, however, that the cooling should be directed to the elimination of self-absorption and to a general improvement of the line shape of the source, rather than to a mere reduction of Doppler width. The atomic beam filtering[4] suggested by Jerrold R. Zacharias should of course lead to a simultaneous elimination of Doppler and collision broadening, as well as of self-absorption. In practice, so far, air or water-cooling have made it possible to obtain interferometric measurements with a signal-to-noise ratio remaining as great as 30:1 at $\pm$ 300 mm path difference with a Hg 198 electrodeless source, the present limit in path difference being set by mechanical limits in carriage motion. There should be no great difficulty in considerably extending this range by adequately improving the effective line shape of the source.

LENGTH INFORMATION IN A DYNAMICAL INTERFEROMETER WITH A PERFECTLY MONOCHROMATIC SOURCE

The use of a split beam or Fabry-Perot interferometer for the measurement of the distance between two stations of a translating carriage is well known. The procedure involves in general two mechanisms. (1) A device for determining when the moving mirror M_1 is at the initial and at the final station respectively. (2) A device for determining the number of wavelengths, or fringes, between the two stations, the number not being an integral in general. The two mechanisms do not necessarily have to be distinct, and the measurement is significant only when the mirror M_1 was stationary first in the initial position, then in the final one. Visual observation is used, measuring the position of a certain fringe portion with respect to a reference mark. In Figure 3 we show a typical arrangement used in dynamical length measurements, providing continuous information. This arrangement is representative of the one used for the translation part of the M.I.T. diffraction grating ruling engine[3]. The interference fringe system in the focal plane of the lens L_2 is an equal inclination circular ring system. Only a small portion of the centre of the ring system is used, as shown in Figure 4.[3] Here we will consider the dynamic portion for the time being: information about the position of M_1 is obtained only during the period when M_1 is moving. It is easily seen that the phase in the interferometer will change by 2π each time M_1 has moved by $\lambda/2$ or 1 fringe, when the interferometer is illuminated by a perfectly

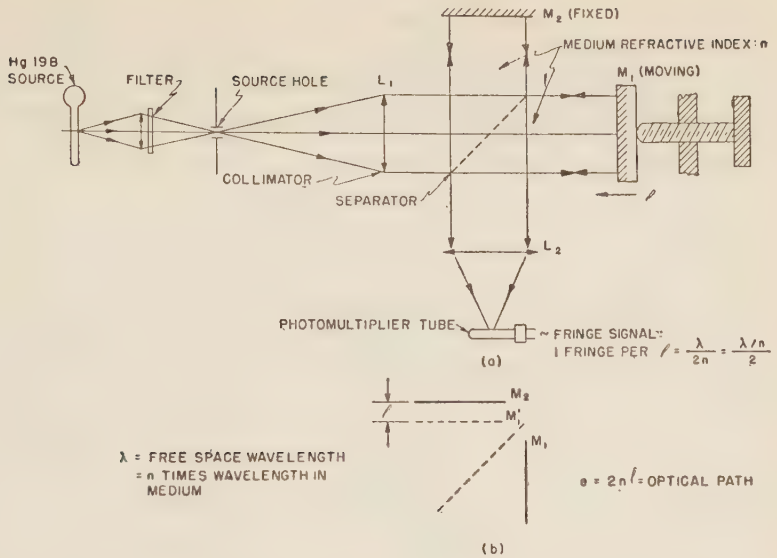


Figure 3

Interferometer arrangement with photoelectric detection. The fringe system in the focal plane of L_2 is an equal inclination ring system. The source hole effectively limits the portion of the ring system used in the interferometer to a small fraction of the central ring for a given mirror separation, as shown in Figure 4. The filter is such as to select the desired wavelength from the single isotope source, and may include such arrangements as are needed to limit the flux conveying information to the natural width of the line. The symmetry of the visibility curve about zero path distance is affected to a great degree by the parallelism of the mirrors M_1 and M_2. Furthermore, the translation has to be normal to the mirror M_1 in order to avoid errors in the effective length per fringe.

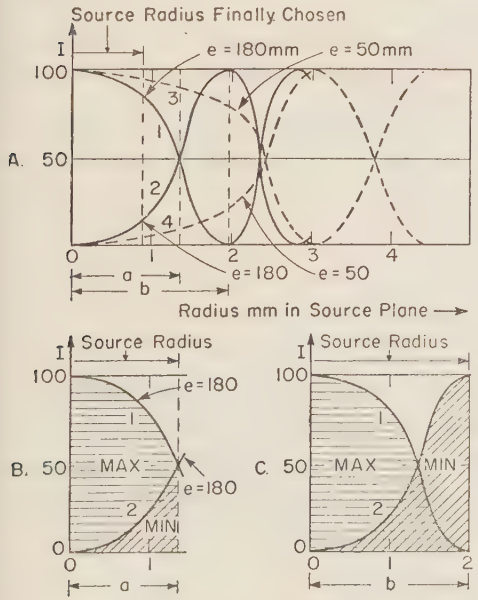


Figure 4

Diagram showing the variation of intensity distribution in interference ring system and effective limitation of flux by source hole. (a) Cross-section through fringe pattern at various mirror separations (Note that the mirror separation is described here as e instead of l as throughout the rest of this article.) (b) Case of mirror separation of 180 mm (path difference 360 mm). (c) Same case as in (b) but with large source hole. The source hole chosen in (a) refers to the M.I.T. ruling engine. For best signal to noise ratio it is necessary to centre the source hole accurately on the ring system which can be achieved by actually observing the ring system in the focal plane of L_2 or its equivalent, and enlarging the source by diffusion. Both the collimator lens and the mirrors M_1 and M_2 must be separately auto-collimated, in that order.

monochromatic collimated beam. The flux falling on the photomultiplier tube is thus qualitatively expected to be modulated in a sinusoidal manner with distance. However, a more accurate investigation of the modulation is found to be necessary.

The intensity distribution in the ring system produced in the focal plane of L_2 by a perfectly monochromatic source is

$$I = 2a^2(1+\cos\varphi), \tag{1}$$

that is

$$I = 4a^2\cos^2(\varphi/2) \tag{2}$$

where a is the amplitude of the monochromatic vibration in each interferometer beam and

$$\varphi = (2\pi e/\lambda)\cos\alpha, \tag{3}$$

α being as shown in Figure 5, λ being the *free space wavelength* and

$$e = 2ln$$

Figure 5
Diagram describing notation used in text.

being the optical path difference between M_1 and M_2 where the medium has a refractive index equal to n. The total flux F falling on the photomultiplier tube is seen to be given by integration over the aperture of the source[3]:

$$F_{\text{MON}} = 2\int_{R=0}^{R_s} I \cdot 2\pi R\, dR,$$

that is

$$F_{\text{MON}} = 16\pi a^2 \int_{R=0}^{R_s} \cos^2(\varphi/2) R\, dR. \tag{4}$$

This integral relation which describes the modulation of light flux F_{MON} in an interferometer using equal inclination fringes and photoelectric detection was first given by Harrison and Stroke[3]. The integration is easily carried out using the following notation:

$$\varphi_0 = 2\pi e/\lambda \ ;$$

We thus have

$$\begin{aligned}\varphi &= \varphi_0 - \varphi_s = 2\pi e/\lambda - (2\pi e/\lambda)\cos\alpha \\ &= (2\pi e/\lambda)(1-\cos\alpha) \\ &\approx (2\pi e/\lambda)\alpha^2/2 \\ &\approx (\pi e/\lambda)R^2/f^2 .\end{aligned}$$

We find that

$$F_{\text{MON}} = 4\pi a^2 R_s \left[1 + \frac{\sin\left[(\pi ln/\lambda)\,\alpha_s^2\right]}{(\pi ln/\lambda)\alpha_s^2} \cos\frac{2\pi ln(1-\alpha_s^2/4)}{\lambda/2} \right]$$

or approximately

$$F_{\text{MON}} = 4\pi a^2 R_s \left[1 + \frac{\sin\left[(\pi ln/\lambda)\,\alpha_s^2\right]}{(\pi ln/\lambda)\,\alpha_s^2} \cos\frac{2\pi l}{[(\lambda/n)/2](1+\alpha_s^2/4)} \right] \tag{5}$$

where a, R_s, λ and α_s are constants determined by the interferometer geometry and the source. The fringe modulation we normally see on the oscilloscopes looking at the photomultiplier current, and shown in Figure 1, is contained in the

$$\cos\frac{2\pi l}{[(\lambda/n)/2](1+\alpha_s^2/4)}$$

factor. It shows the important point that with a finite source aperture α_s the observed fringe spacing is actually obtained in terms of a longer effective wavelength

$$\lambda_{\text{eff}} \approx \lambda(1+\alpha_s^2/4)$$

Alternatively it can be said that the *effective length per fringe* is $[(\lambda/n)/2](1-\alpha_s^2/4)$ approximately. Physically the correction arises from the fact that the mirror separation is not only measured normally for a normal translation, but also at slight angles up to $\pm\,\alpha_s$, and the phototube registers the integrated effect. Figure 6 shows the classical picture of the path difference for a beam passing through the interferometer at a angle α. We get one fringe each time

$$l = (\lambda/n)/2\cos\alpha,$$

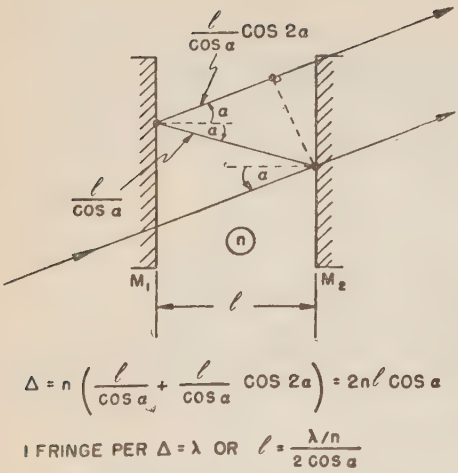


Figure 6
Diagram showing path difference corresponding to beam traversing mirrors at an angle. This case can also arise if mirrors do not move normally to a beam which is well auto-collimated in collimator lens alone.

that is to say that one has to move the mirror more than $(\lambda/n)/2$ to get one fringe when a beam at an angle α is used. The correction factor $(1-\alpha_s^2/4)$ can be calculated with all the needed precision. It is shown here only as resulting from the second term in the expansion of a cosine. This correction factor is not negligible. For a source hole of 1.8 mm in diameter and $f = 1500$ mm

$$\alpha_s \approx 0.9/1500 \cong 6/10^4$$
$$\alpha_s^2/4 \approx 1/10^7.$$

For a length measurement of say $l = 136$ mm approximately 5×10^5 fringes of $\lambda = 5461$ Å, the length correction is

$$5\times10^5/10^7 = 5/100 \text{ fringe} = 5 \text{ centrifringes.}$$

This factor has therefore to be taken into account when absolute length measurements in terms of a given light wave are being performed, such as in the case of measurements of scales and etalons. Another typical case where this situation arises is shown in Figure 7.[5] This figure shows a schematic diagram of an apparatus that is now being built at M.I.T. for a determination of the speed of light to 1 part in 10^8. In this measurement we need a precision of 1 centifringe in $\pm$ 5 inches, or an even greater one for measurements over 1 inch or so. We wish to note that an otherwise accurate length measurement might be in error due to a misinterpretation of the effective length measured.

LENGTH INFORMATION IN A DYNAMICAL INTERFEROMETER WITH A DOPPLER BROADENED SOURCE

The next situation which deserves investigation is the one resulting from a broadening of the source. This case has considerable practical importance as far as the "visibility"

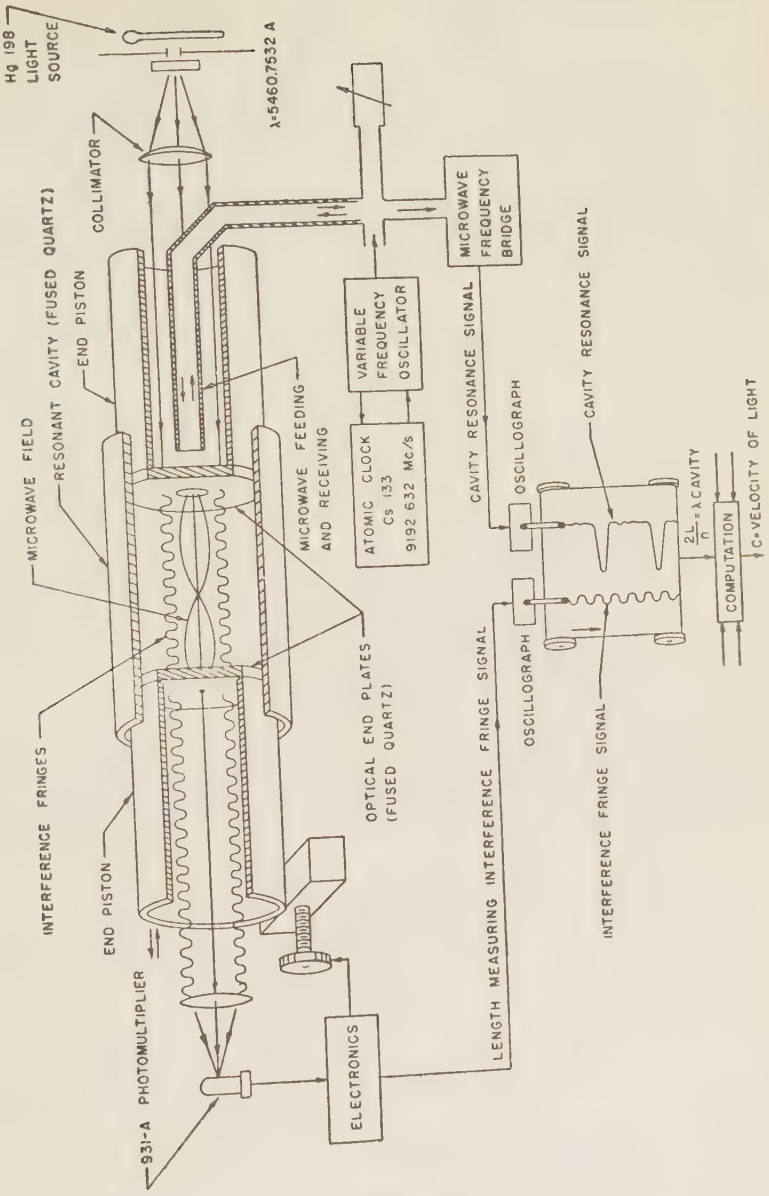


Figure 7

Speed of light measurement in terms of two primary standards, Cs 133 and Hg 198. This diagram illustrates a typical application where the desired accuracy of 10^{-8} makes it imperative to take into account the various fine points discussed in the present article. The M.I.T. ruling engine will be used for the translation of the end-pistons in the resonant cavity.

and information content of the fringe signal at great mirror separations are concerned. Candler[6] lists several of the well known causes of line broadening. Among these the Doppler effect, collision broadening, pressure broadening, self-reversal and broadening due to a foreign gas seem to be accessible to empirical reduction, while this is not immediately possible for the "natural" width. The broadening due to a foreign gas might lead to asymmetric line shapes, and may also play a not negligible role in self-absorption.

It is possible to calculate the effect of an exponential broadening, such as the Doppler broadening, on the fringe visibility law obtained in eq. (5) above. Let us write the Doppler broadened intensity distribution as

$$I_\nu = I_0 \exp\left[-(mc^2/2kT)(\nu-\nu_0)/\nu_0)^2\right]$$

or

$$I_\nu = I_0 e^{-(\Delta\nu)^2/\sigma^2}$$

where

$m = MM_H$
M = atomic weight of gas
M_H = mass of hydrogen atom = 1.668×10^{-24} gm
$k = 1.380 \times 10^{-16}$ erg °K^{-1}
$c \simeq 2.998 \times 10^{-10}$ cm sec^{-1}
T = absolute temperature.

ν and ν_0 are given in wave-number units (cm^{-1}). For the flux modulation F_D on the phototube in the case of a Doppler broadened source and the interferometer examined above, we can write eq. (5) in the form

$$F_D = 4\pi R_s^2 a^2 \int_{-\infty}^{+\infty} \left\{ 1 + \frac{\sin\left[(\pi ln/\lambda)a_s^2\right]}{(\pi ln/\lambda)a_s^2} \cos\left[(2\pi e/\lambda)(1-a_s^2/4)\right] \right\} e^{-(\Delta\nu)^2/\delta^2} d\nu$$

where

$$e = 2ln$$

or approximately

$$F_D = 4\pi R_s^2 I_0 \int_{-\infty}^{+\infty} \left[1 + \frac{\sin(\pi la_s^2 n)\nu}{(\pi la_s^2 n)\nu} \cos 2\pi e\nu \right] e^{-(\Delta\nu)^2/\sigma^2} d\nu. \qquad (6)$$

In order to reduce the integration to standard forms we write F_D in the form

$$F_D = A\int_{-\infty}^{+\infty} \left[1 + \frac{\sin a(\nu_0+\Delta\nu)}{a\nu_0}\cos b(\nu_0+\Delta\nu)\right] e^{-(\Delta\nu)^2/\sigma^2}\, d\nu$$

$$= A\int_{-\infty}^{+\infty} \left[1 + \frac{\sin a(\nu_0+q)}{a\nu_0}\cos b(\nu_0+q)\right] e^{-q^2/\sigma^2}\, dq.$$

We will consider $a\nu_0$ to be constant in the range of integration. After trigonometric transformation we get

$$F_D = A\int_{-\infty}^{+\infty} e^{-\alpha^2 q^2}\, dq + (A/2a\nu_0)\int_{-\infty}^{+\infty} \sin\left[s(q+r/s)\right] e^{-\alpha^2 q^2} dq$$

$$+ (A/2a\nu)\int_{-\infty}^{+\infty} \sin\left[t(q+\varphi/t)\right] e^{-\alpha^2 q^2}\, dq$$

where $\alpha^2 = 1/\sigma^2$. We find that the integral

$$\int_{-\infty}^{+\infty} e^{-\alpha^2 x^2} \sin\left[p(x+\lambda)\right] dx = (\sqrt{\pi}/\alpha)\, e^{-p^2/4\alpha} \sin p\lambda$$

is listed in Bierens de Haan (269/2). Using this we find

$$F_D = (A\sqrt{\pi}/\alpha)\left\{1 + (2/a\nu_0)\left[e^{-(a+b)^2\sigma^2/4}\sin(a+b)\nu_0 + e^{-(a-b)^2\sigma^2/4}\sin(a-b)\nu_0\right]\right\}$$

where $a = \pi e a_s^2/2$ and $b = 2\pi e$. We next try to simplify the terms in the square brackets. These are of the form

$$S = e^{-m}\sin(a\nu_0 + b\nu_0) + e^{-n}\sin(a\nu_0 - b\nu_0)$$

giving

$$S = (e^{-m}+e^{-n})\sin a\nu_0 \cos b\nu_0 + (e^{-m}-e^{-n})\cos a\nu_0 \sin b\nu_0.$$

Noting the orders of magnitude of m and n, we can write

$$e^{-m} + e^{-n} \approx 2e^{-m} = 2\exp\left[-(\pi e)^2\sigma^2\right]$$

and

$$e^{-m} - e^{-n} \approx -\sigma^2 ab = -2\sigma^2(\pi e)^2 a_s^2;$$

Finally, with these simplifications, and reintroducing $a\nu_0$ into the brackets, and noting again that $e = 2ln$, we get

$$F_D = 4\pi^{3/2} R_s^2 a^2 \sigma \left\{ 1 + e^{-4\pi^2 l^2 \sigma^2 n^2} \frac{\sin(\pi l a_s^2 n/\lambda)}{(\pi l a_s^2 n/\lambda)} \cos\left[\frac{2\pi l(1-a_s^2/4)n}{\lambda/2} \right] \right.$$

$$\left. - 2\sigma^2 (\pi l n)^2 a_s^2 \cos(\pi l a_s^2 n/\lambda) \sin\left[\frac{2\pi l n(1-a_s^2/4)}{\lambda/2} \right] \right\}:$$

Numerical values again show that the third term in the brackets is small in practice.

The final expression for the fringe modulation in the case of a Doppler broadened source can be written as

$$F_D = 4\pi^{3/2} R_s^2 a^2 \sigma \left\{ 1 + e^{-4\pi^2 l^2 \sigma^2 n^2} \frac{\sin(\pi l a_s^2 n/\lambda)}{(\pi l a_s^2 n/\lambda)} \cos\left[\frac{2\pi l}{[(\lambda/n)/2]\ (1+a_s^2/4)} \right] \right\}: \quad (7)$$

This equation reduces to the form of eq. (5) for small values of l where the Doppler broadening should have little effect. We note the important point that *the exponential term is important*. For example, for $\lambda = 5461$ Å, $T = 300°\mathrm{K}$, the exponent of e is approximately equal to

$$-4l^2/10^5$$

with l in mm and $n \approx 1$. The integral of eq. (7) was calculated by us independently for the general case of photoelectric reception. It appeared subsequently that Rayleigh[2] had performed a somewhat similar calculation for the case of visual reception. While it is not immediately obvious that the modulation factor of the cosine factor in eq. (7) and the corresponding visibility factor of Rayleigh should have similar forms, this appears to be true.

Examination of the orders of magnitude involved in the visibility modulation factors of eq. (7) shows that the exponential factor predominates over the $(\sin x)/x$ factor for mirror separation ranges up to say 500 mm or so.

The effect of a source broadening of this nature is:

(a) A decrease in signal amplitude with path distance.
(b) A decrease in signal to noise ratio with path distance.
(c) A very slight phase shift in the fringe phase at great path distances.

A simple calculation shows that with a Hg 198 source of 1.9 effective Doppler widths, at a temperature of 300°K the phase shift due to the exponential factor $\exp[-4\pi^2 l^2 \sigma^2 n^2]$ will be 3×10^{-7} fringes for a mirror separation $l = 114.22$ mm (order of interference 4×10^5), and $n \approx 1$.

EXPERIMENTAL RESULTS AND COMPARISON WITH THEORY

After numerical values had been obtained for eq. (7) we examined once more the experimental curve shown in Figure 2, obtained with an uncooled Hg 198 source,

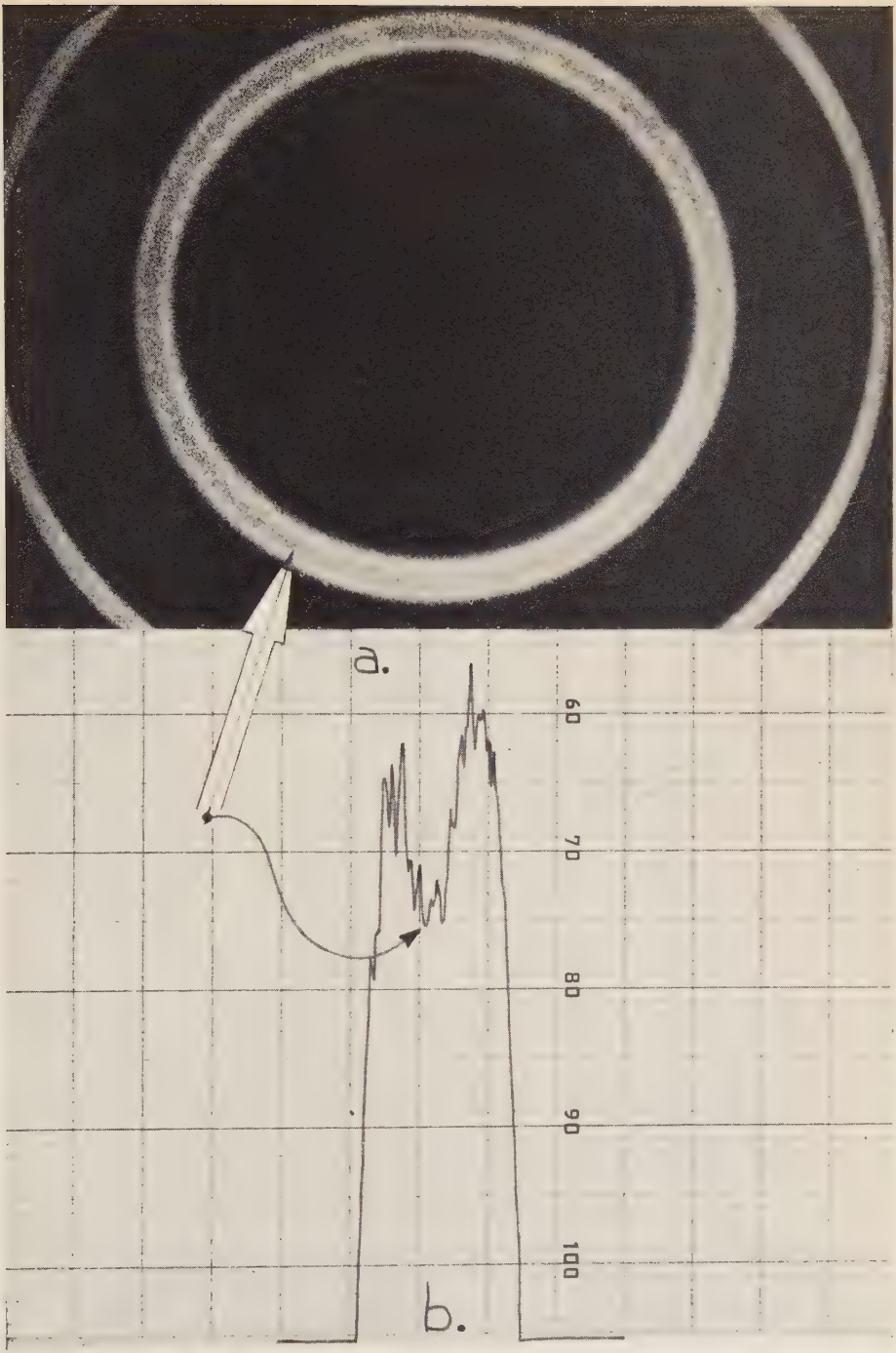


Figure 8

Fabry-Perot interferogram and microphotometer trace through first ring corresponding to self-reversed Hg 198 source. (a) Shows how the line was self-reversed in this discharge tube containing about 0.001 gm of Hg 198 at about 40°C, the tube having 5 mm bore, and being viewed across its diameter. (b) Shows a microphotometer trace made on the negative. This self-reversed line gave rise to the visibility curve shown in Figure 2.

excited with a 200 mc/s oscillator in an electrodeless discharge at about 30 watts. The tube was a Meggers[7] tube of 6 mm OD, 5 mm ID, about 130 mm long, made of Vycor. It is said to contain about 0.001 gm of the mercury isotope Hg 198 and argon at about 5 mm (Hg) pressure. Under operating conditions the tube temperature is about 40°C. The minima at $\pm$ 100 mm and the secondary maxima at $\pm$ 140 mm of the curve in Figure 2 did not correspond to the ones resulting from the $(\sin x)/x$ factor. It was quickly suspected that a second source wavelength near the main wavelength, and of almost the same intensity, could be responsible for this visibility curve shape. The Fabry-Perot interferogram of the source under the above mentioned operating conditions is shown in Figure 8. It shows that the 5461 Å line used was strongly self-reversed and could indeed account for the experimental curve obtained in Figure 2. The minima correspond, with the most simplifying assumptions, to a wavelength difference of about 5×10^{-6} between the two maxima of the self-reversed line.

Several visibility curves with both air cooling and water cooling were then measured. These are shown in Figure 9. The ordinates represent again the amplitude of the fringe signal as appearing on the oscilloscope (Figure 1). These amplitudes are also represented in the $[\sin(\pi e a_s^2/\lambda)/(\pi e a_s^2/\lambda)] \exp[-4\pi^2 l^2 \sigma^2 n^2]$ factor. A moderately strong air blast was used, and very little difference was found to exist between air-cooled and water-cooled tubes, the latter having a slightly better signal to noise ratio at

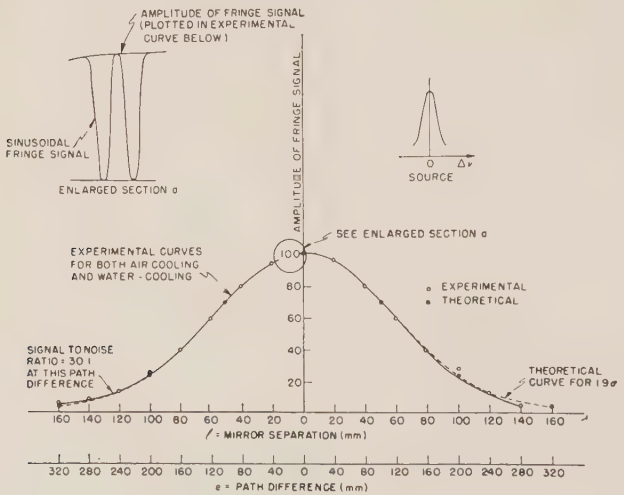


Figure 9

Experimental visibility curves corresponding to cooled source conditions, and comparison with theoretical curve for 1.9σ showing very good accord. The effective line shape of the source can be determined from such visibility curves by Fourier analysis in the general case. In this figure, the fit to an exponential line shape is very good. The effective width of 1.9σ, greater than the one corresponding to the Doppler width parameter σ under the given experimental conditions, is attributed to broadening by self-absorption.

path differences of the order of 300 mm. No self-reversal was observed in the Fabry-Perot interferometer with 2.5 cm spacers under these cooling conditions. The actual cooling conditions are difficult to describe, but it was estimated that the Vycor glass tube was kept at around 20°C. We note that the experimental curves of Figure 9 do not show any secondary maxima in the measurement range. Indeed, they can be fitted very closely with a theoretical exponential visibility curve of the form shown in eq. (7). However, instead of corresponding to a "pure" Doppler source of the form $\exp[-(\Delta\nu)^2/\sigma^2]$, they appear to result from a source of the form $\exp[-(\Delta\nu)^2/3.6\sigma^2]$. A very good fit to the experimental curves corresponding to eq. (7) is obtained with this source form assumption. Great interest was naturally attached to the explanation of such a curious source form. It had the same general exponential shape as a Doppler broadened source, but was about twice as wide at the half-width.

Since we had strong experimental evidence showing that self-absorption in the source was sufficiently great to lead to observable self-reversal with the uncooled source at only slightly higher temperatures, it was normal to inquire whether a smaller residual amount of self-absorption would not lead to a source form $\exp[-(\Delta\nu)^2/3.6\sigma^2]$ which seemed to result from actual experiments.

MODIFICATION OF DOPPLER BROADENED SOURCE BY SELF-ABSORPTION AND CORRESPONDING VISIBILITY CURVES

In order to simplify the estimate of the source shape, we have made the following assumptions:

(a) The source is Doppler broadened of the form

$$F_0 = e^{-(\Delta\nu)^2/\sigma^2};$$

(b) The source emission layer is near the centre of the tube.

(c) The absorption is according to an exponential law with distance, leading to an effective source of the form

$$F_\nu = F\, e^{-K\nu x}$$

K_ν being the absorption coefficient, x being the distance from the centre to the walls of the tube.

(d) The absorption coefficient is assumed to be of the form

$$K_\nu = \beta e^{-(\Delta\nu)^2/\sigma}$$

with $\sigma^2 \approx 10^{-4}$ at 300°K. In mercury at this temperature, the Doppler half-width $2\Delta\nu_D = 10^{-6}\nu_0$ approximately. Therefore, up to the half-width, we have $(\Delta\nu)^2/\sigma^2 \approx 0.8$. At one half the half-width $\Delta\nu \approx 4.5\times10^{-3}$, and $(\Delta\nu)^2/\sigma^2 \approx 0.2$. In the central region of the source we can thus easily use the first two terms of the series expansion of exp $[-(\Delta\nu)^2\sigma/^2]$. We obtain for the self-absorbed Doppler broadened source the form

$$F_\nu \approx e^{-(1-\beta x)(\Delta\nu)^2/\sigma^2} e^{-\beta x} . \tag{8}$$

For a greater fraction of the source width it is possible to include the contribution of the higher terms of the series expansion by properly assigning a numerical coefficient α to βx inside the brackets of the exponential. This leads to the form

$$F_\nu \simeq e^{-(1-\alpha\beta x)(\Delta\nu)^2/\sigma^2} e^{-\beta x} . \tag{8'}$$

From the experimental values obtained in Figure 9 we should therefore have

$$1 - \alpha\beta x = 1/3.6$$

or

$$\alpha\beta x = 0.72 :$$

Assuming $x = 0.25$ cm, we obtain $\beta = K_0 \approx 3\alpha$, where K_0 is the absorption coefficient for the central wave-number ν_0, and $\alpha = 1$ for $\Delta\nu < 0.005$.

We have thus shown that a sufficiently great remainder of the self-absorption, which we have experimentally observed to be very large in the "uncooled" sources at about 40°C, could easily yield a source form that could account for the effective broadening by a factor of 1.9 of the Doppler broadened source as resulting from interferometric visibility measurements.

As to the interpretation of the absorption coefficient K_0, there is sufficient evidence in the literature to point to the complexity of the estimate[8-12]. The vapour pressure of mercury at 40°C is 0.006079 and at 20°C it is 0.001201 mm (Hg). These values alone are not sufficient to account for the observed self-reversal at 40°C. The presence of argon and the nature of the high-frequency discharge make it desirable to investigate further their contribution to the broadening, some aspects of which have been previously studied[8-12].

CONCLUSION

For precise interferometric length measurements over great distances two conditions are necessary:

(a) Precise knowledge of the instrumental laws determining the fringe signal.

(b) Good "visibility", that is good signal to noise ratio over the measurement range.

These two points have been investigated for the case of photoelectric detection. It was shown that practical source apertures could lead to appreciable errors in the effective length measured. Source conditions encountered in a Hg 198 discharge tube excited by electrodeless discharge, may result in self-absorption of a sufficiently great magnitude to broaden effectively the 5461 Å line by a factor of the order of two as compared to its Doppler width at the temperature of about 20°C. While it is known that cooling of the source from room temperature to about — 180°C would be necessary to reduce the Doppler width by a factor of two, proper source conditions, even at room temperature, should reduce the effective source line-width by this same factor (that is to its Doppler width), provided that the self-absorption is adequately reduced. The atomic beam filtering suggested by Zacharias[4] should allow to obtain effective source widths of the order of the natural widths and appears therefore to be the promising way of attaining the needed excellent visibility over very large path distances. The use of sources or lines with smaller broadening should be considered. We have found in this series of experiments with a Michelson Twyman-Green interferometer that a water-cooled Hg 198 electrodeless discharge gave the very good signal to noise ratio of 30:1 at path differences as large as $\pm$ 320 mm, both air and water cooling giving comparable results. Very good accord between theoretical and measured visibility curves is found with the assumption of residual self-absorption based on experimental evidence. This accord permits the use of the theoretically derived results for estimation of fringe phase shifts introduced by instrumental conditions in the interferometric system, and allows to arrange these conditions so as to minimize systematic measurement errors. Adequately small source apertures and sources of small effective line breadth are found to be sufficient to assure a predictable accuracy in interferometric length measurements with photoelectric detection.

ACKNOWLEDGMENTS

Professor Lee C. Bradley III has kept a most stimulating interest in the problems described above and has independently calculated equations (5) and (7). The author wishes to thank him for his active contribution in discussions and private communications. We plan to publish results of further experiments and theory in a joint paper. The author also wishes to thank Dean George R. Harrison and Professor Jerrold R. Zacharias who have encouraged this study and contributed ideas and suggestions.

This work was supported in part by the Army (Signal Corps), the Air Force (Office of Scientific Research, Air Research and Development Command), the Navy (Office of Naval Research) and the Office of Ordnance Research (Contract DA 19-020-ORD-1887).

REFERENCES

1. Michelson, A. A., 1927, *Studies in Optics*, University of Chicago Press, p. 34.
2. Lord Rayleigh (J. W. Strutt), 1915, *Phil. Mag.*, **29**, 274.
3. Harrison, G. R. and Stroke, G. W., 1955, *J. opt. Soc. Amer.*, **45**, 112.

4. Zacharias, J. R., M.I.T., RLE Quart. Rep., Jan. 15, 1956, p. 68.
5. Zacharias, J. R., Harrison, G. R., Stroke, G. W., Mason, S. J. and Searle, C. L., *ibid.*, p. 59.
6. Candler, C., 1951, *Modern Interferometers*, 9—49.
7. Meggers, W. F. and Westfall, F. O., 1950, *J. Res. nat. Bur. Stand.*, **44,** 447.
8. Nutting, P. G., 1906, *Astrophys. J.*, **23,** 220.
9. Kuech, R. and Retschinsky, T., *Ann. Phys.*, **20,** 563· *ibid.*, 1907, **22,** 595.
10. Metcalfe, E. P. and Venkatesachar, B., 1922, *Proc. roy. Soc.*, **A 100,** 149.
11. Slater, J. C., 1925, *Phys. Rev.*, **25,** 783.
12. Harrison, G. R. and Slater, J. C., 1925, *ibid.*, **26,** 176.

DISPOSITIFS INTERFERENTIELS POUR L'OBSERVATION DES OBJETS TRANSPARENTS

M. Francon

Professeur à la Sorbonne, et à l'Institut d'Optique, Paris

SOMMAIRE

On décrit quelques dispositifs interférentiels à polarisation utilisables pour l'observation des objets transparents isotropes. Ces dispositifs ont un champ d'observation étendu et sont particulièrement adaptés aux études dans le domaine macroscopique.

Depuis quelques années le développement des méthodes d'observation des objets transparents a été considérable, d'abord avec le contraste de phase puis avec les procédés interférentiels. Parmi ces derniers, il faut mentionner spécialement les appareils basés sur la polarisation car ils permettent de réaliser des montages très simples.

Considérons (Figure 1) un objet transparent A d'indice de réfraction n situé à l'intérieur d'un milieu transparent L d'indice n' limité par des faces planes et parallèles. C'est par example un défaut d'homogénéité à l'intérieur d'une lame de verre à faces parallèles. Un rayon lumineux incident tel que 1 qui a traversé l'objet A n'a pas effectué le même chemin optique que le rayon 2 qui passe à côté. Si la surface d'onde Σ incidente est plane, après traversée de l'objet, elle devient Σ' et est déformée par suite des variations de chemin optique.

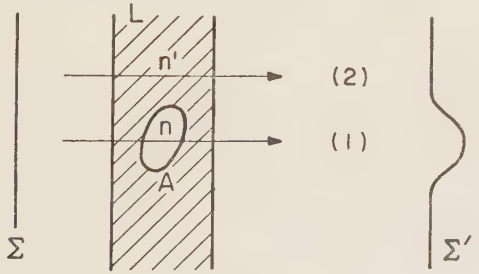


Figure 1

Supposons que l'on place après L un système biréfringent Q quelconque entre un polariseur $\mathcal{P}$ et un analyseur $\mathcal{A}$ (Figure 2). Soit en A' l'image de l'objet A donnée par un système optique quelconque non représenté sur la Figure 2. Par suite de la présence du biréfringent Q l'onde Σ' est dédoublée dans le plan de

Received March 23, 1956.

l'image A' en 2 ondes O et E. Ces 2 ondes O et E ont un décalage latéral d (Figure 3) qui dépend des caractéristiques du biréfringent Q et un décalage en profondeur Δ qui représente la différence de marche entre l'onde O et l'onde E. Cette différence de marche Δ dépend non seulement des caractéristiques du biréfringent Q mais aussi de la façon dont il est utilisé comme nous le préciserons plus loin.

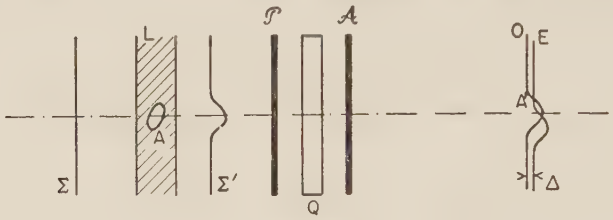


Figure 2

Grâce au polariseur et à l'analyseur ces 2 ondes O et E peuvent interférer suivant les règles classiques des interférences en lumière polarisée. Si la différence de marche est Δ dans les régions planes 1 et 2 elle sera $\Delta \pm \alpha d$ dans les régions où la pente de la surface d'onde (O ou E) sera $\pm \alpha$. Supposons que par un réglage convenable de l'orientation et de la position du biréfringent Q nous ayons la teinte sensible

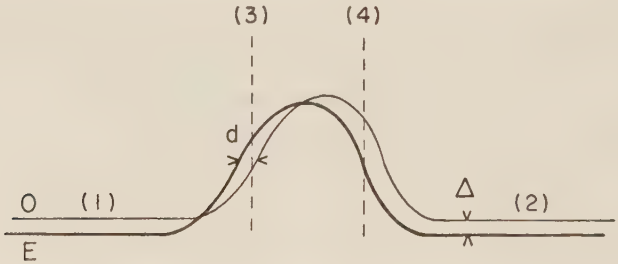


Figure 3

pourpre dans les régions 1 et 2. Dans les régions 3 et 4 la teinte sensible va virer et l'on pourra ainsi déceler les variations de chemin optique produites par l'objet A. Il est ainsi possible de rendre visible un objet transparent comme par la méthode du contraste de phase mais il y a une différence que nous précisons maintenant.

En écrivant la relation $\Delta \pm \alpha d$ nous avons supposé implicitement que le décalage latéral d est petit par rapport à la largeur de l'objet A à étudier. La méthode permet ainsi de déceler les pentes de la surfaces d'onde, c'est-à-dire les variations du chemin optique et non le chemin optique lui-même. On a une méthode différentielle. Il y a donc une différence essentielle avec la méthode du contraste de phase qui donne le chemin optique et non sa dérivée. Toutefois si le décalage latéral d est plus grand que l'objet (Figure 4) chaque image interfère avec le fond de l'autre et la méthode donne alors le chemin optique lui-même cette fois exactement comme dans la méthode du contraste de phase.

Précisons maintenant les caractéristiques du système biréfringent Q. De nombreux dispositifs ont été proposés qui ont leurs avantages et leurs inconvénients. Ceux que nous proposons s'appliquent spécialement à l'étude des objets macroscopiques ou alors aux faibles grossissements.

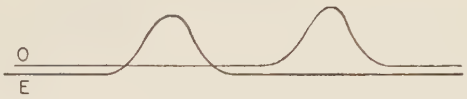


Figure 4

Le dédoublement de l'onde incidente en deux ondes émergentes est produit par un dispositif analogue au prisme biréfringent bien connu de Wollaston. Dans ce dernier prisme la séparation angulaire est due à un ensemble de deux prismes d'un cristal uniaxe (quartz par exemple) et dont les axes sont croisés (Figure 5). Si le faisceau indicent étroit passe par le centre du prisme où les deux éléments

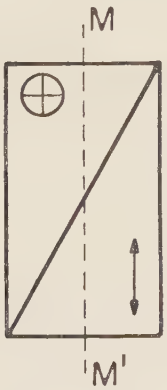


Figure 5

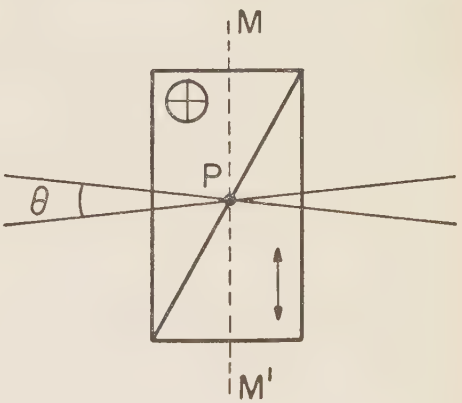


Figure 6

ont même épaisseur, les deux rayons O et E sortent en phase. Ceci se traduit sur les Figures 2 et 3 par $\Delta = 0$. Si le faisceau étroit se déplace le long de la ligne intérieure MM' la différence de marche Δ entre O et E varie à volonté. Dans les montages qui seront décrits plus loin, on forme l'image de la source lumineuse sur MM' à l'intérieure du prisme (Figure 6). Dans ces conditions le prisme est traversé par un faisceau lumineux d'angle θ qui n'est pas très petit ou alors le champ d'observation devient trop faible. Tout se passe dans la région P comme si on avait deux lames de quartz à faces parallèles et croisés. On sait que dans ce cas on observe à l'infini des hyperboles équilatères. Le champ ne peut être uniforme qu'au centre de ces hyperboles et dans une petite région. Dans le but d'augmenter le champ, nous avons remplacé les lames de quartz par deux lames taillées l'une dans un cristal positif (quartz) l'autre taillée dans un cristal négatif (spath) et dont les épaisseurs sont choisies convenablement.

Soient i l'angle d'incidence sur une lame cristalline uniaxe (Figure 7), e son épaisseur, $n_0 - n_e$ sa biréfringence. La différence de marche Δ entre les deux rayons émergents (non représentés sur la Figure 7) dépend de l'orientation de l'axe du cristal.

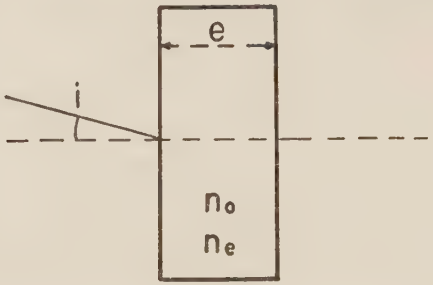


Figure 7

Supposons la lame taillée parallèlement à l'axe. Si le plan d'incidence contient l'axe optique on a

$$\Delta = e(n_0 - n_e)\,[1 - (i^2/2n_0^2)]. \tag{1}$$

Dans le cas où le plan d'incidence est perpendiculaire à l'axe on a

$$\Delta = e(n_0 - n_e)\,[1 + (i^2/2n_0 n_e)]. \tag{2}$$

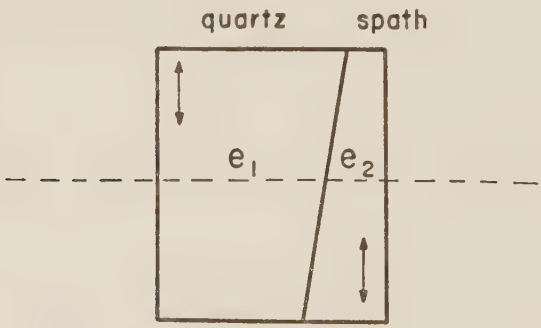


Figure 8

Réalisons alors la combinaison représentée par la Figure 8. Pour le plan d'incidence perpendiculaire à l'axe optique les retards Δ_1 et Δ_2 à la traversée des prismes de quartz et de spath s'écrivent:

$$\Delta_1 = e_1(n_0 - n_e)\,[1 + (i^2/2n_0 n_e)] \qquad \text{(quartz)}$$

$$\Delta_2 = e_2(n_0 - n'_e)\,[1 + (i^2/2n'_0 n'_e)]; \qquad \text{(spath)}$$

choisissons les épaisseurs e_1 et e_2 au centre du prisme de façon que

$$e_1 |(n_0 - n_e)| = e_2 |(n'_0 - n'_e)|\ ;$$

la différence de marche totale à la traversée de l'ensemble s'écrit en remarquant que pour le quartz $n_0 - n_e < 0$ et pour le spath $n'_0 - n'_e > 0$

$$\Delta = e_2(n'_0 - n'_e) \frac{n_0 n_e - n'_0 n'_e}{2 n_0 n_e n'_0 n'_e} i^2, \tag{3}$$

et pour le plan d'incidence contenant l'axe

$$\Delta'' = e_2(n_0 - n'_e) \frac{n'^2_0 - n_0^2}{2 n_0^2 n'^2_0} i^2. \tag{4}$$

Les courbes isochromatiques du système sont encore des hyperboles mais ces hyperboles sont plus larges.

Dans le cas d'un prisme de Wollaston ordinaire on aurait:

$$\Delta = e(n_0 - n_e) \frac{n_0 + n_e}{2 n_0^2 n_e} i^2,$$

soit pour le quartz:

$$\Delta = e(n_0 - n_e) \times 0.42\, i^2. \tag{5}$$

Pour l'ensemble quartz-spath la relation (3) donne:

$$\Delta' = e_2 (n'_0 - n'_e) \times 0.006\, i^2$$

et la relation (4)

$$\Delta'' = e_2(n'_0 - n'_e) \times 0.028\, i^2.$$

Par comparaison avec la relation (5), on voit que pour une même valeur de $e(n_0 - n_e)$ on atteint la même différence de marche pour un angle 8 fois plus grand dans le cas Δ' et 4 fois plus grand dans le cas Δ''. Le champ de ce prisme est donc 4 fois plus grand que celui du Wollaston ordinaire.

Une deuxième solution donnant un champ considérable est représentée par la Figure 9.

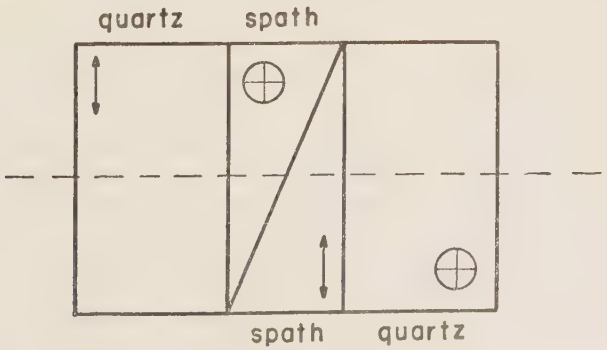


Figure 9

Lorsque le plan d'incidence est perpendiculaire à l'axe du premier spath d'entrée, on aura pour ce spath:

$$\delta' = e'(n'_0 - n'_e)\ [1 + (i^2/2n'_0\, n'_e)], \tag{6}$$

et pour le quartz qui suit:

$$\delta = e(n_0 - n_e)[1 - (i^2/2n_0^2)]\ ; \tag{7}$$

à la traversée de ces 2 premiers prismes d'axes croisès, on aura la différence de marche $\Delta' = \delta' - \delta$

$$\Delta' = [e'(n'_0 - n'_e) + e(n_e - n_0)]\left[1 + \frac{[e'(n'_0 - n'_e)/2n'_0 n'_e] - [e(n_e - n_0)/2n_0^2]}{e'(n'_0 - n'_e) + e(n_e - n_0)} i^2\right]. \tag{8}$$

Pour le plan d'incidence contenant l'axe du premier spath on a à la traversèe de ce premier spath:

$$\delta' = e'(n'_0 - n_e)\ [1 - (i^2/2n'^2_0)], \tag{9}$$

et pour le quartz

$$\delta = e(n_0 - n_e)[1 + (i^2/2n_0 n_e)] \tag{10}$$

à la traversée des deux:

$$\Delta'' = [e'(n'_0 - n'_e) + e(n_e - n_0)]\left[1 + \frac{[e(n_e - n_0)/2n_0 n_e] - [e'(n'_0 - n'_e)/2n'^2_0]}{e'(n'_0 - n'_e) + e(n_e - n_0)} i^2\right]. \tag{11}$$

On aura des lignes isochromatiques circulaires si les 2 coefficients de i dans les expressions (8) et (11) sont égaux, c'est-à-dire

$$\frac{e}{e'} = \frac{n'^2_0 - n'^2_e}{n_e^2 - n_0^2} \times \frac{n_0^2\, n_e}{n'^2_0\, n'_e} = 17.331. \tag{12}$$

En remplaçant dans (8) ou (11) le rapport e/e' tiré de (12) on a:

$$\Delta = [e'(n'_0 - n'_e) + e(n_e - n_0)]\ (1 + 0.005\, i^2). \tag{13}$$

Pour le système quartz-spath placé à la suite on a une différence de marche égale en valeur absolue mais de signe contraire à (13). Ceci bien entendu pour le centre du prisme où les épaisseurs de quartz sont égales. Donc la différence de marche totale après traversée du système dans quatre lames est ègale à zéro même pour des rayons fortement inclinés et on a un champ considérable. Ceci est vrai quelle que soit la longueur d'onde grâce à la faible différence de dispersion du quartz et du spath. Le système peut donc aussi être utilisé en lumière blanche sans être gêné par le chromatisme.

On peut réaliser différents montages avec ces prismes; en voici deux qui sont parmi les plus intéressants.

Sur la Figure 10 l'objet est en P contre un miroir sphérique M_2. Au moyen d'un objectif O_1 et après réflexion sur un miroir M_1 on forme en S' une image de la source S. L'image S' est au voisinage du centre du miroir M_2 qui en donne une nouvelle image en S''. Le prisme est en Q sur les images S' et S''. Un objectif O_2 donne de P une image en P' qui est projetée sur un écran E ou observée à l'oculaire. Le polariseur $\mathcal{P}$ est placé avant Q et l'analyseur $\mathcal{A}$ après Q sur le faisceau de retour. Grâce à la double traversée du prisme Q les faisceaux se compensent à l'aller et au retour et il n'est pas nécessaire de diaphragmer le système.

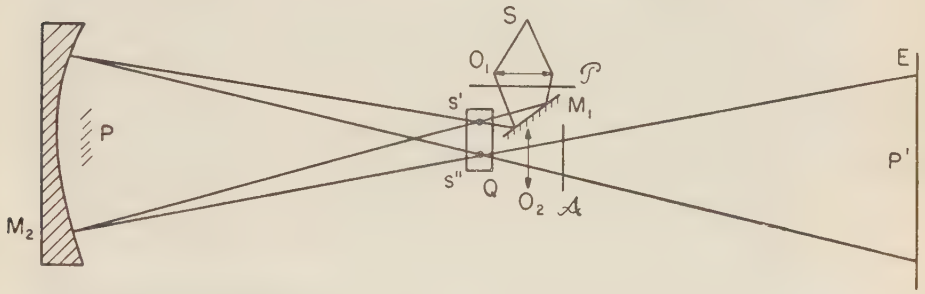


Figure 10

La Figure 11 montre un montage analogue mais où l'objet est traversé seulement une fois.

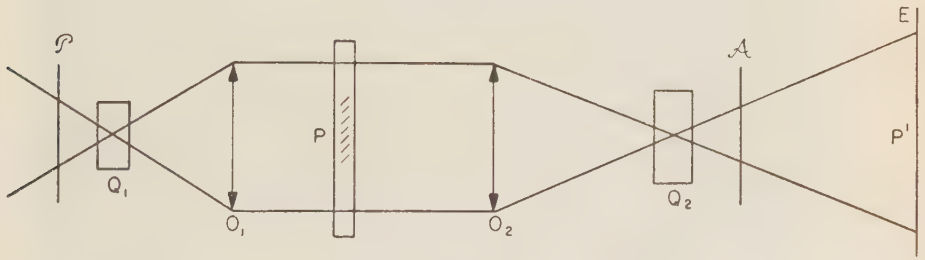


Figure 11

L'objet en P est placé entre 2 objectifs O_1 et O_2. Pour éviter toute diaphragmation on conjugue 2 prismes Q_1 et Q_2 identiques mais croisés.

Les photographies 12 et 13 montrent les résultats obtenues avec la méthode différentielle c'est-à-dire lorsque le décalage latéral d est petit. L'objet est ici constitué par 2 liquides d'indices différents ne formant pas un mélange homogène. Les liquides sont contenus dans une cuve à faces parallèles et les variations de chemin optique se traduisent par des variations de teintes comme il a été dit précèdemment. Connaissant le dédoublement d et ayant repéré les teintes par comparaison avec les teintes de l'echelle de Newton, on peut mesurer les variations de chemin optique.

Si on rapproche ou si on éloigne Q du miroir M_2 on observe un système de franges sur l'écran E et les déformations de ces franges permettent également de mesurer les variations du chemin optique. Pour obtenir le même effet sur la Figure 11, il suffit de déplacer l'un ou l'autre des prismes le long de l'axe.

Les prismes dont nous venons de donner la description permettent de réaliser des montages simples et lumineux. Il est ainsi possible d'observer et d'étudier des objets isotropes transparents caractérisés par des variations d'indice ou d'épaisseur soit par une méthode différentielle à faible dédoublement soit par une méthode à dédoublement total donnant directement le chemin optique mais applicable dans le cas de petits objets isolés pour qu'il n'y ait pas enchevêtrement des images. La sensibilité de la méthode différentielle dépend essentiellement du dédoublement choisi et des dimensions de l'objet. Dans le cas de la méthode du dèdoublement total la sensibilité peut atteindre très facilement 1/200 de la longueur d'onde sur la mesure du chemin optique.

Variations du chemin optique à l'interieur d'une masse non homogène de liquides

Figure 12. Réglage de l'interféromètre sur le pourpre du 1er ordre de l'échelle des teintes de Newton ($\Delta = 0.565$ microns)

Figure 13. Réglage de l'interféromètre sur le bleu du 2ème ordre ($\Delta = 0.600$ microns)

SOME FINE MECHANICAL INSTRUMENTS PRODUCED IN ISRAEL

E. Alexander
Department of Physics, The Hebrew University of Jerusalem

As in most other countries, the fine mechanics industry in Israel is carried on in small enterprises with a great variety of size, production programme and function.

Several fine mechanical products of the Israel industry are described below.

kWh-meters — Elco (*Figure* 1)

About 6,000—7,000 units are produced per month. The stringent accuracy requirements demanded by the Palestine Electric Corporation guarantee a high quality product. The special feature of this meter is a construction which allows high overloads without serious loss in accuracy.

Watermeter — Arad (*Figure* 2)

This meter is manufactured at Kibbutz Dalia. It represents a typical example of successful house industry in a communal settlement in the field of fine mechanics, a field which should have a number of advantages for this purpose.

The meter is manufactured under licence and supervision of Swiss Aquameter, Basel. About 70 workers, mostly members of Dalia, are employed in the production of these meters. About 3,000 meters of different sizes ($\frac{1}{2}''$ to $8''$) are produced monthly. The overall accuracy is $\pm 2\%$.

Ammeter — Sekely-Hoffmann (*Figure* 3)

These ammeters and voltmeters are produced with an accuracy of $\pm 2\%$ for switch board instruments and $\pm 1\%$ for laboratory instruments. Mainly A.C. meters are produced, but D.C. instruments are also manufactured in smaller quantities.

Alarm clock-work — Orlogin (*Figure* 4)

The clock-work reproduced in Figure 3 is produced by Orlogin under a Swiss licence, but with a revised construction. It is the only clock-work produced in its entirety in Israel.

Compass — Goldberg Instruments (*Figure* 5)*; Refractometer — Goldberg Instruments* (*Figure* 6)

Prof. Goldberg's plant has until now been the only one in Israel manufacturing several types of scientific instruments of original design, and is therefore, according

Received August 31, 1956.

to a more limited definition, the only representative of a fine mechanical and optical industry in Israel.

It should be stressed that the examples given are not intended to give a comprehensive picture either of the industry as a whole (since a number of not less important plants have not been mentioned), nor of the full production programmes of the plants mentioned, which are generally more extensive.

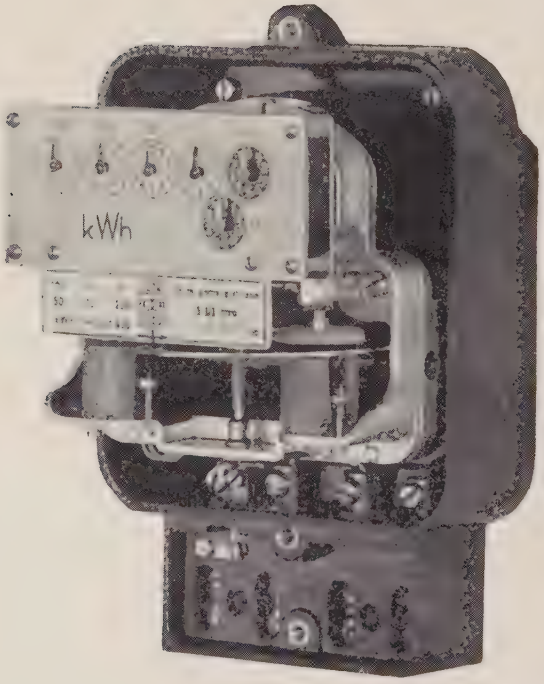


Figure 1

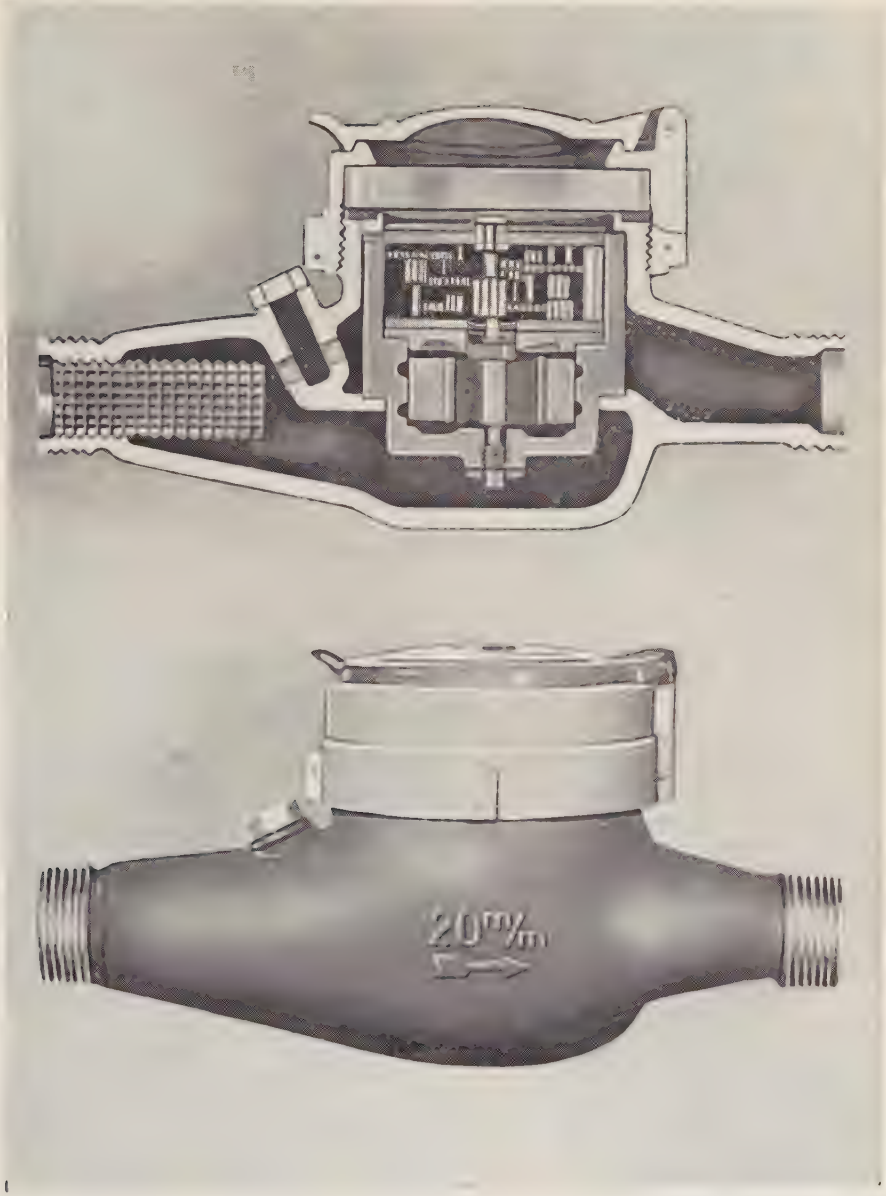


Figure 2

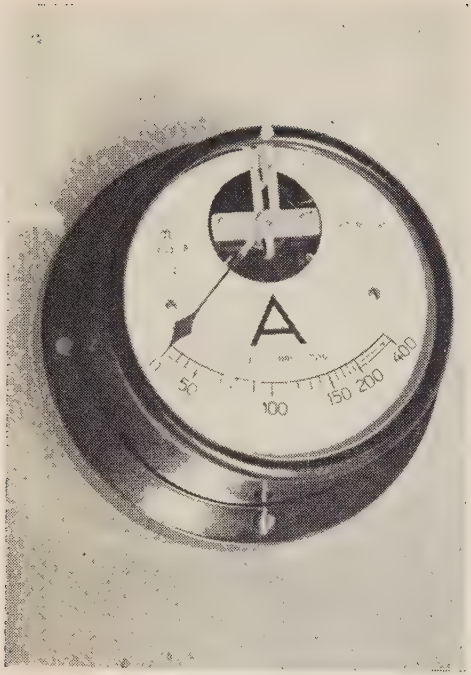


Figure 3

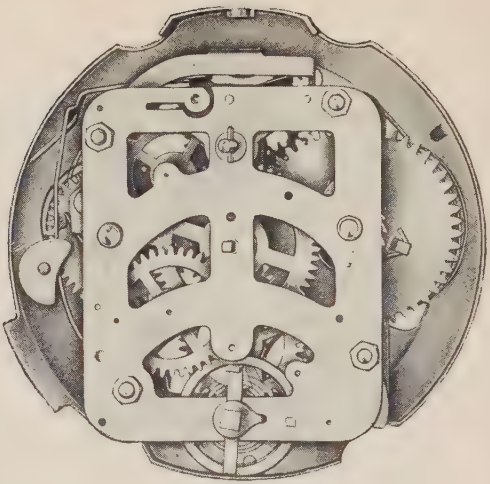


Figure 4

Figure 5

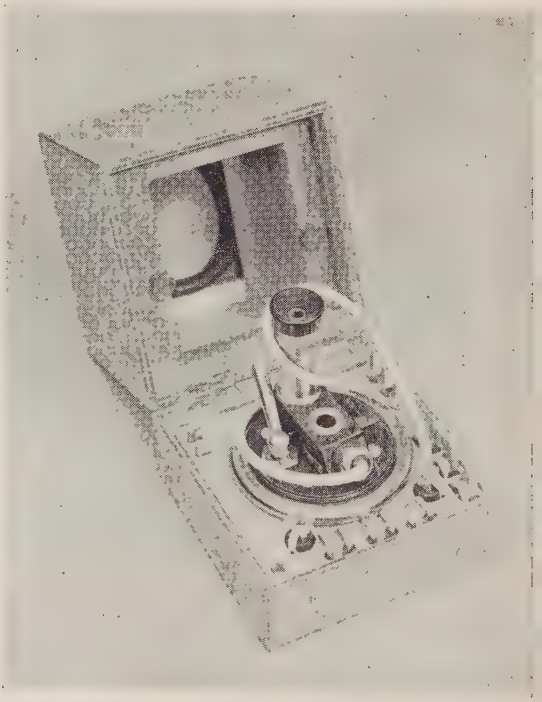


Figure 6

Volume 5C, 1955—1957

BULLETIN OF THE RESEARCH COUNCIL OF ISRAEL

Section C TECHNOLOGY

Bull. Res. Counc. of Israel. C. Techn.

Incorporating the Scientific Publications of the Technion — Israel Institute of Technology, Haifa

INDEX

TO

VOLUME 5C

The Weizmann Science Press of Israel . Jerusalem

BULLETIN OF THE RESEARCH COUNCIL OF ISRAEL

יוצא לאור ע"י

מוסד ויצמן לפרסומים במדעי חטבע ובטכנולוגיה בישראל

המועצה המדעית לישראל • משרד החנוך והתרבות • האוניברסיטה העברית בירושלים

הטכניון — מכון טכנולוגי לישראל • מכון ויצמן למדע • מוסד ביאליק

Published by

THE WEIZMANN SCIENCE PRESS OF ISRAEL

Research Council of Israel • Ministry of Education and Culture

The Hebrew University of Jerusalem • Technion—Israel Institute of Technology

The Weizmann Institute of Science • Bialik Institute

Manuscripts should be addressed: The Editor, The Weizmann Science Press of Israel, P.O.B. 801, Jerusalem
33, King George Ave., Telephone 62844

מ.י.ם 2157/1127—6.57/2000

CONTENTS

Number 1, December 1955

Number 2—3, March—June 1956

Number 4, June 1957

AUTHOR INDEX

SUBJECT INDEX

NOTICE TO CONTRIBUTORS

Contributors to the *Bulletin of the Research Council of Israel* should conform to the following recommendations of the editors of this journal in preparing manuscripts for the press.

Contributions must be original and should not have been published previously. When a paper has been accepted for publication, the author(s) may not publish it elsewhere unless permission is received from the Editor of this journal.

Papers may be submitted in English, French and Russian.

MANUSCRIPT

General

Papers should be written as concisely as possible. MSS should be typewritten on one side only and double-spaced, with side margins not less than 2.5 cm wide. Pages, including those containing illustrations, references or tables, should be numbered.

The Editor reserves the right to return a MS to the author for retyping or any alterations. Authors should retain copies of their MS.

Spelling

Spelling should be based on the Oxford Dictionary and should be consistent throughout the paper. Geographic and proper names in particular should be checked for approved forms of spelling or transliteration.

Indications

Greek letters should be indicated in a legend preceding the MS, as well as by a pencil note in the margin on first appearance in the text.

When there is any room for confusion of symbols, they should be carefully differentiated, e.g. the letter "l" and the figure "1"; "O" and "0".

Abbreviations

Titles of journals should be abbreviated according to the *World List of Scientific Periodicals.*

Abstract

Every paper must be accompanied by a brief but comprehensive abstract. Although the length of the abstract is left to the discretion of the author, 3% of the total length of the paper is suggested.

References

In Sections A and C, and in Letters to the Editor in all Sections, references are to be cited in the text by number, e. g., ... Taylor[3] ..., and are to be arranged in the order of appearance.

In Sections B and D, the references are to be cited in the text by the author's name and date of publication in parenthesis, e.g., ... (Taylor 1932).... If the author's name is already mentioned in the text, then the year only appears in the parenthesis, e. g., ... found by Taylor (1932)... . The references in these Sections are to be arranged in alphabetical order.

The following form should be used :

3. TAYLOR, G. I., 1932, *Proc. roy. Soc.*, **A138**, 41.

Book references should be prepared according to the following form:

4. JACKSON, F., 1930, *Thermodynamics*, 4th ed., Wiley, New York.

TYPOGRAPHY

In all matters of typography the form adopted in this issue should be followed. Particular attention should be given position (of symbols, headings, etc.) and type specification.

ILLUSTRATIONS

Illustrations should be sent in a state suitable for direct photographic reproduction. Line drawings should be drawn in large scale with India ink on white drawing paper, bristol board, tracing paper, blue linen, or blue-lined graph paper. If the lettering cannot be drawn neatly by the author, he should indicate it in pencil for the guidance of the draftsman. Possible photographic reduction should be carefully considered when lettering and in other details.

Half tone photographs should be on glossy contrast paper.

Illustrations should be mounted on separate sheets of paper on which the caption and figure number is typed. Each drawing and photograph should be identified on the back with the author's name and figure number.

The place in which the figure is to appear should be indicated in the margin of the MS.

PROOFS

Authors making revisions in proofs will be required to bear the costs thereof. Proofs should be returned to the Editor within 24 hours, otherwise no responsibility is assumed for the corrections of the author.

REPRINTS

Reprints may be ordered at the time the first proof is returned. A table designating the cost of reprints may be obtained on request.

Orders in America should be addressed to Interscience Publishers Inc., New York, N. Y., and in England and Europe to Wm. Dawson & Sons, Ltd., Cannon House, Macklin Street, London, W. C. 2, directly or through booksellers.

Annual subscription per section (four issues): IL.4.000 ($5.50, £ 2)
Single copy IL.1.000 ($1.50, 12 s.)

PRINTED BY GOVERNMENT PRESS, JERUSALEM

SET ON MONOTYPE